



## DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

<b>(51) Classification internationale des brevets <sup>6</sup> :</b> <b>C12N 15/82, 15/15, C07K 14/81, C12N 5/10, A01H 5/00, 5/10, C07K 16/38, A61K 38/57, 7/00</b>	<b>A1</b>	<b>(11) Numéro de publication internationale: WO 99/38987</b> <b>(43) Date de publication internationale: 5 août 1999 (05.08.99)</b>
<b>(21) Numéro de la demande internationale: PCT/FR99/00195</b> <b>(22) Date de dépôt international: 29 janvier 1999 (29.01.99)</b> <b>(30) Données relatives à la priorité:</b> 98/01089                      30 janvier 1998 (30.01.98)                      FR <b>(71) Déposant (pour tous les Etats désignés sauf US):</b> MERISTEM THERAPEUTICS [FR/FR]; 8, rue des Frères Lumière, F-63100 Clermont-Ferrand (FR). <b>(72) Inventeurs; et</b> <b>(75) Inventeurs/Déposants (US seulement):</b> GRUBER, Véronique [FR/FR]; Résidence Sainte-Madeleine, 44 C, avenue Jean-Jaurès, F-63400 Chamalières (FR). OLAGNIER, Béatrice [FR/FR]; 6, rue de Wailly, F-63000 Clermont-Ferrand (FR). BOURNAT, Philippe [FR/FR]; 48, rue Bonnabaud, F-63000 Clermont-Ferrand (FR). THEISEN, Manfred [DE/FR]; 78, boulevard F. Mitterrand, Bâtiment K, F-63000 Clermont-Ferrand (FR). MEROT, Bertrand [FR/FR]; 6, place du Communal, La Coussedière, F-63530 Volvic (FR).		<b>(74) Mandataires:</b> GUTMANN, Ernest etc.; Ernest Gutmann - Yves Plasseraud S.A., 3, rue Chauveau-Lagarde, F-75008 Paris (FR). <b>(81) Etats désignés:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, brevet ARIPO (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Publiée</b> <i>Avec rapport de recherche internationale.</i> <i>Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues.</i>
<b>(54) Title:</b> METHOD FOR PRODUCING, BY PLANT CELLS, $\alpha_1$ -ANTITRYPSIN AND ITS ALLELES, AND PRODUCTS CONTAINING THE RESULTING $\alpha_1$ -ANTITRYPSIN <b>(54) Titre:</b> PROCEDE DE PRODUCTION, PAR DES CELLULES VEGETALES, D' $\alpha_1$ -ANTITRYPSINE ET DE SES VARIANTES, ET PRODUITS CONTENANT L' $\alpha_1$ -ANTITRYPSINE AINSI OBTENUE <b>(57) Abstract</b> <p>he invention concerns a method for producing <math>\alpha_1</math>-antitrypsin (<math>\alpha_1</math>-AT), or alleles thereof characterised in that it consists in: (i) introducing in the monocotyledon plant cells, a nucleic acid comprising a sequence coding for <math>\alpha_1</math>-antitrypsin (<math>\alpha_1</math>-AT) or for an allele thereof, the alleles being differentiated from natural human <math>\alpha_1</math>-AT by one or several substitution(s), deletion(s) or insertion(s) of amino acids, and optionally a sequence coding for a selective agent; (ii) selecting the cells having integrated the nucleic acid in stable manner; (iii) propagating the transformed cells in culture, or regeneration of entire chimerical or transgenic plants; (iv) recuperating and optionally purifying the resulting <math>\alpha_1</math>-antitrypsin or its alleles.</p> <b>(57) Abrégé</b> <p>L'invention concerne un procédé de production d'<math>\alpha_1</math>-antitrypsine (<math>\alpha_1</math>-AT), ou de variantes de celle-ci caractérisé par: i) l'introduction dans des cellules végétales monocotylédones, d'un acide nucléique comportant une séquence codant pour l'<math>\alpha_1</math>-antitrypsine (<math>\alpha_1</math>-AT) ou pour une variante de celle-ci, les variantes se différenciant de l'<math>\alpha_1</math>-AT humaine naturelle par une ou plusieurs substitution(s), délétion(s) ou insertion(s) d'acides aminés, et éventuellement une séquence codant pour un agent sélectif; ii) sélection des cellules ayant intégré de manière stable l'acide nucléique; iii) propagation des cellules transformées en culture, ou régénération de plantes entières chimériques ou transgéniques; iv) récupération et éventuellement purification de l'<math>\alpha_1</math>-antitrypsine ou de ses variantes, ainsi produites.</p>		

### UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	B Brésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

PROCEDE DE PRODUCTION, PAR DES CELLULES VEGETALES,  
D' $\alpha_1$ -ANTITRYPSINE ET DE SES VARIANTES, ET PRODUITS  
CONTENANT L' $\alpha_1$ -ANTITRYPSINE AINSI OBTENUE

L'invention concerne un procédé de production, par des cellules végétales monocotylédones, d' $\alpha_1$ -antitrypsine et de ses analogues, et les protéines ainsi obtenues. L'invention vise également les cellules et plantes monocotylédones, génétiquement transformées, capables de produire des analogues d' $\alpha_1$ -antitrypsine, et les construits d'acides nucléiques impliqués dans la transformation. En outre, l'invention vise des produits pharmaceutiques contenant l' $\alpha_1$ -antitrypsine et ses variantes, ainsi obtenues.

L'alpha-1-antitrypsine,  $\alpha_1$ -AT, également connue sous les désignations d'alpha-1-inhibiteur de protéase ( $\alpha_1$ -PI), d'antitrypsine et de metserpine, est une protéine plasmatique présente dans le plasma à une concentration de l'ordre de 1,3 grammes par litre. Synthétisée par le foie, il s'agit d'une glycoprotéine de 51kDa à 54kDa, constituée d'une seule chaîne polypeptidique de 394 acides aminés (Long et al., Biochemistry, 1984, vol. 23, pages 4828-4837). Elle possède 3 sites de N-glycosylation sur les résidus Asn-46, Asn-83 et Asn-247, les trois chaînes glycosidiques ayant une structure "core" commune de Asn-(NAc Glc)<sub>2</sub>-(Man)<sub>3</sub>, avec 2 ou 3 antennes de NAc Glc-Gal-Sia. La partie glucidique représente 13% de la masse totale de la glycoprotéine.

En raison d'un polymorphisme survenant à 3 positions dans la protéine, il existe 3 formes d' $\alpha_1$ -AT:

M1, M2, M3 :

	213	376	101
M1	Val/Ala	Glu	Arg
M2	Val	Asp	Arg
M3	Val	Asp	His

L' $\alpha_1$ -AT est un inhibiteur de protéase à sérine. Sa fonction biologique principale est d'inhiber l'élastase neutrophile et de protéger ainsi le tissu pulmonaire contre les attaques protéolytiques.

Il existe plusieurs variantes "pathologiques" communes de la protéine : par exemple la forme "Z" dans laquelle Glu-342 est substitué par Lys, et la forme "S" dans laquelle Glu-264 est substitué par Val. Des différences au niveau de la glycosylation, et en particulier dans le nombre de résidus d'acide sialique, peuvent aussi être observées.

Les mutants Z et S, bien que capables de fonctionner comme agents anti-élastase, sont produits en faible quantité conduisant à une déficience, et surtout dans le cas de la variante "Z", au développement d'emphysème pulmonaire sévère.

Cette pathologie est actuellement traitée par thérapie substitutive avec des concentrés plasmatiques. Les besoins annuels actuels par patient sont estimés à 200 grammes et la demande mondiale à plusieurs centaines de kilogrammes par an. L'échelle de production de la protéine à partir de sang humain couvre difficilement ces besoins.

La production d' $\alpha_1$ -AT humaine par voie recombinante a donc été envisagée. L' $\alpha_1$ -AT recombinante non-glycosylée a notamment été produite chez E. coli sous forme de protéine de fusion. La protéine,

exprimée à un taux d'environ 15% des protéines cellulaires totales, présentait une activité anti-élastase caractéristique d' $\alpha_1$ -AT, mais était selon les auteurs, susceptible de présenter des propriétés immunologiques modifiées en raison de l'extension N-terminale (Courtney et al., PNAS, 1984, vol. 81, pages 669-673).

La production chez la levure d' $\alpha_1$ -AT à des taux élevés (10-3% des protéines cellulaires totales) a également été décrite par Johansen et al. (Mol. Biol. Med., 1987, vol. 4, pages 291-305). Le rendement a été augmenté par la délétion des premiers 5, 10 ou 15 acides aminés.

Travis et al. (J. Biol. Chem., 1985, vol. 260 n° 7, pages 4384-4389) ont décrit la production d' $\alpha_1$ -AT non glycosylée chez la levure. Deux variantes de la protéine ont été obtenues à un rendement d'environ 10% des protéines cellulaires solubles totales, chacune ayant une activité  $\alpha_1$ -AT mais une stabilité réduite par rapport à l' $\alpha_1$ -AT native.

L' $\alpha_1$ -AT glycosylée a été exprimée dans le sang et dans le lait d'animaux transgéniques, notamment des lapins, des souris et des brebis.

Dans le sang des animaux transgéniques, la protéine était présente à une concentration d'environ 1 mg/ml de plasma. Elle conserve son activité anti-élastase (Massoud et al., C. R. Acad. Sci. Paris, 1990, tome 311, série III, pages 275-280).

Dans le lait de brebis transgéniques, les niveaux d' $\alpha_1$ -antitrypsine ont atteint 60 g/l. La protéine ainsi produite est entièrement glycosylée (Wright et al., Bio/Technology, Septembre 1991, vol. 9, pages 830-834). La nature de la glycosylation n'a pas été publiée.

La production de protéines recombinantes dans le lait des animaux transgéniques se propose de diminuer

les coûts de production. Cependant, il reste des problèmes d'éthique et de contaminations virales et subvirales du même type que ceux rencontrés avec les produits purifiés classiques.

Le problème technique que se propose de résoudre la présente invention est de produire de l' $\alpha_1$ -AT recombinante en grande quantité à des coûts faibles, sans risque de contaminations virales ou sub-virales, par exemple par des prions. La protéine présente une activité  $\alpha_1$ -AT satisfaisante, et doit être stable et adaptée du point de vue immunologique à la forme d'administration choisie. Les présents inventeurs ont résolu ce problème technique en utilisant, comme hôte pour la transformation génétique, des cellules végétales, et plus particulièrement des cellules végétales monocotylédones.

Diverses équipes se sont déjà intéressées à la production de protéines recombinantes de mammifères dans des cellules végétales ou dans des plantes transgéniques. Par exemple, l'expression spécifique dans les graines de colza, de la leu-enképhaline a été obtenue avec des niveaux d'expression d'environ 0,1% (Vanderkerckhove et al., Biotechnology, 1989, vol. 7, pages 929-932).

En 1990, Simons et al. (Biotechnology, 1990, vol. 8, pages 217-221) ont transféré le gène de la sérum albumine humaine dans des cellules de tabac et de pomme de terre. Quelle que soit l'origine des peptides signaux (humaine ou végétale), des taux de sérum albumine humaine de l'ordre de 0,02% des protéines totales ont été obtenus dans les feuilles, les tiges et les tubercules de pomme de terre.

D'autres protéines recombinantes de mammifères ont aussi été produites dans les plantes : l'antigène de surface de l'hépatite B (Mason et al., P.N.A.S., 1992, vol. 89, pages 11745-11749) ; l'interféron

humain (Edelbaum, J. of Interferon Res., 1992, vol. 12, pages 449-453) ; un anticorps de souris anti-Streptococcus mutans, agent de la carie dentaire (Hiatt et Ma, FEBS, 1992, vol. 307, pages 71-75) ; un anticorps anti-Herpès et l'hirudine (Russel, Moloney, respectivement, Intl. Meeting of Production of Recombinant Proteins in Plants, Leicester, 1994, page 43, pages 36-38).

L'ensemble de ces recherches montre que la production de protéines recombinantes de mammifères dans les cellules végétales est possible et que les mécanismes de synthèse de protéines à partir des séquences d'ADN sont similaires chez les cellules animales et les cellules végétales. De plus, il apparaît que la cellule végétale est capable de réaliser la plupart des étapes de maturation post-traductionnelles des protéines comme l'assemblage des sous-unités, le repliement des chaînes, la maturation protéolytique, la formation des ponts-disulfures et le clivage des peptides signaux de manière semblable aux cellules animales.

Quelques différences existent néanmoins au niveau de la glycosylation. Bien que les étapes de glycosylation conduisant aux glycanes polymannosidiques soient réalisées de manière identique chez les plantes et chez les mammifères, la maturation des glycanes polymannosidiques en glycanes complexes est sensiblement différente. Les N-glycanes des glycoprotéines végétales diffèrent principalement des glycanes de mammifères par l'absence d'acide sialique, également connu sous le nom d'acide N-acétylneuraminique, et par la présence d'un résidu de  $\beta$ -1,2 xylose et d'un résidu de fucose lié en  $\alpha$ -1,3 au résidu de GlcNAc proximal du "core" (Driouich et al., Regard sur la biochimie, 1993, vol. 3, pages 33-42). Les abréviations signifiant des

résidus glucidiques sont celles habituellement utilisées dans l'art (VLIEGENTHART J.F.G and MONTREUIL J., Chap. II, Glycoproteins, Montreuil J., Schachter H., Vliegenthart J.F.G., 1995, ELSEVIER SCIENCE B.V., pages 13-28 ; MONTREUIL J., Advances in Carbohydrate Chemistry and Biochemistry, 1980, Vol. 37, Academic Press Inc., pages 157-223).

La glycosylation joue un rôle important dans la mise en place de la structure spatiale de la protéine, dans la solubilité de la molécule, dans la protection contre les attaques protéolytiques et dans les mécanismes de reconnaissance immunitaire. Le glycanne contrôle aussi la demi-vie de la protéine et peut lui-même, dans certains cas, porter la fonction biologique. Compte tenu du rôle important joué par la glycosylation, l'on essaie normalement, pour la fabrication de protéines thérapeutiques, de reproduire le plus fidèlement possible la glycosylation naturelle de la protéine. Ceci n'est pas possible dans des cellules végétales, en raison de la présence du xylose, absent chez les animaux, et de l'absence d'acide sialique chez les plantes. Bien que présentant de nombreux avantages du point de vue économique, les différences au niveau de la glycosylation de la protéine peuvent donc rendre imprévisible le succès de l'utilisation de cellules végétales pour la production de protéines thérapeutiques.

Malgré ces différences importantes, les présents inventeurs ont réussi à produire dans des plantes monocotylédones transgéniques, à des niveaux d'expression élevés, des protéines ayant une activité d' $\alpha_1$ -AT. Les protéines ont une stabilité acceptable. De manière surprenante, les inventeurs ont aussi constaté, que dans certaines conditions, la maturation protéolytique de la glycoprotéine est fonction de l'adressage cellulaire, permettant la production d'une



série de "variantes" protéiques, de poids moléculaire différent, présentant toutes une activité d' $\alpha_1$ -AT.

L'invention concerne donc un procédé de production d' $\alpha_1$ -antitrypsine, ou de variantes de celle-ci caractérisé par :

- i) l'introduction, dans des cellules végétales provenant d'une espèce monocotylédone, d'un acide nucléique comportant une séquence codant pour l' $\alpha_1$ -antitrypsine ou pour une variante de celle-ci, les variantes se différenciant de l' $\alpha_1$ -AT humaine naturelle par une ou plusieurs substitution(s), délétion(s) ou insertion(s) d'acides aminés, et éventuellement une séquence codant pour un agent sélectif ;
- ii) sélection des cellules ayant intégré de manière stable l'acide nucléique ;
- iii) propagation des cellules transformées en culture, ou régénération de plantes entières chimériques ou transgéniques ;
- iv) récupération et éventuellement purification de l' $\alpha_1$ -antitrypsine ou de ses variantes, ainsi produites.

L'invention concerne également une protéine susceptible d'être produite par le sus-dit procédé, et présentant une activité de protéase à sérine, de préférence une activité de l' $\alpha_1$ -AT, caractérisée en ce que :

- soit elle est exempte de chaînes N-glycosidiques ;

- soit elle est N-glycosylée par au moins un glycanne de type mannosidique ou polymannosidique, et/ou par au moins un glycanne de type complexe comportant éventuellement au sein de sa structure un ou plusieurs résidus de xylose et éventuellement un ou plusieurs résidus de fucose, le glycanne de type complexe étant normalement dépourvu de résidus d'acide sialique.

Dans le contexte de l'invention, les termes ci-après ont les significations suivantes :

- une "variante" de l' $\alpha_1$ -AT signifie une protéine qui se différencie de l' $\alpha_1$ -AT humaine naturelle (mature ou pré-protéine) par une ou plusieurs substitution(s), délétion(s) ou insertion(s) d'acides aminés. De manière générale, les variantes présentent au moins 90% et de préférence au moins 95% d'homologie ou d'identité avec la séquence en acides aminés décrite par Long et al. (Biochemistry, 1984, vol. 23, pages 4828-4837), illustrée dans la Figure 1, le pourcentage d'homologie entre deux protéines étant défini tel qu'indiqué ci-dessous. Le terme « variante » est utilisé, dans le contexte de l'invention, de manière synonyme avec le terme « analogue ». Les variantes présentent aussi une activité d'inhibiteur de protéase à sérine. Il a été constaté que certaines modifications au niveau de la séquence en acides aminés de l' $\alpha_1$ -AT peut conférer à la protéine une activité d'inhibiteur de protéase à sérine autre que celle normalement associée à l' $\alpha_1$ -AT, par exemple une activité caractéristique de l'antithrombine III (Jallat et al., Protein Engineering, Vol. 1, N° 1, pp 29-35, 1986). Les variantes d' $\alpha_1$ -AT de l'invention présentent une glycosylation identique ou différente de l' $\alpha_1$ -AT humaine naturelle. Le terme « variante » englobe également des fragments de l' $\alpha_1$ -AT et ses analogues qui présentent une activité d'inhibiteur de protéase à sérine, de préférence une activité d' $\alpha_1$ -AT.

- « une activité d'inhibiteur de protéase à sérine » signifie la capacité d'inhiber partiellement ou totalement toute protéase ayant un résidu sérine au sein de son site actif, par exemple la trypsine, l'élastine, la cathepsine G, la thrombine, la collagénase, la chymotrypsine et la plasmine.

- "une activité de l' $\alpha_1$ -AT" signifie une activité anti-élastase, déterminée selon la technique de Travis et Johnson, Meth. in Enzymol., 1981, vol. 80, pages 754-765. Cette technique est détaillée dans les exemples. L'activité peut également être déterminée selon la capacité de la protéine à inhiber la trypsine (voir Travis et Johnson, supra). De manière générale, pour les deux techniques, la protéine recombinante présente une activité d'inhibition d'au moins 30% et de préférence au moins 75%, par exemple 85% de celle de l' $\alpha_1$ -AT humaine naturelle à molarité égale. Une troisième technique consiste en la détermination de la capacité de la protéine recombinante à hydrolyser le substrat Methoxy-L-Alanyl-L-Prolyl-L-Valine-P-nitroanilide (Massoud et al., supra). La capacité de la protéine à inhiber l'activité de chacune des protéases à sérine mentionnées ci-dessus peut aussi être utilisée comme indice de l'activité des protéines de l'invention (The Plasma Proteins", Second Edition, Ed. Pulnam, Academic Press, 1975, p.241-242).

- « monocotylédone » signifie un végétal phanérogame angiosperme dont la tige et la racine sont (presque toujours) dépourvues de cambiums et donc de formations secondaires; la feuille est presque toujours pourvue de nervures parallèles; la graine renferme un embryon à un seul cotylédon.

- le pourcentage d'homologie entre deux séquences d'acides aminés est calculé comme étant le nombre d'acides aminés identiques plus le nombre d'acides aminés similaires dans l'alignement des deux séquences, divisé par la longueur des séquences entre deux positions données. Si, entre les deux positions données, les deux séquences n'ont pas la même longueur, le pourcentage d'homologie est le nombre d'acides aminés identiques et similaires, divisé par la longueur de la séquence la plus longue. Les acides

aminés considérés comme étant "similaires" sont connus dans l'art, voir par exemple R.F. Feng, M.S. Jobson and R.F. Doolittle ; J. Mol. Evol. ; 1985 ; vol. 21 ; pages 112-115. Ils sont normalement considérés comme étant ceux qui, au sein d'une matrice de permutation, ont un coefficient de substitution positif.

Selon une première variante de l'invention, la protéine recombinante à activité  $\alpha_1$ -AT est non-glycosylée. Ce type de protéine peut être obtenu dans une cellule végétale en utilisant la séquence d'acide nucléique codant pour le polypeptide mature, c'est-à-dire sans le peptide signal N-terminal de 24 acides aminés (voir par exemple Figure 1). Aucun autre signal n'est ajouté. Le polypeptide ainsi produit s'accumule dans le cytoplasme de la cellule. Il aura normalement un poids moléculaire inférieur d'environ 10kD à celui de l' $\alpha_1$ -AT humaine naturelle, c'est-à-dire autour de 45kDa, et une stabilité moindre dans le plasma.

Selon une variante préférée, la protéine de l'invention est glycosylée à l'un au moins des 3 sites naturels de glycosylation : Asn-46, Asn-83 ou Asn-247. La glycosylation est de type mannosidique, polymannosidique ou de type complexe, ou un mélange des deux. L'expression « entièrement glycosylée » signifie, dans le contexte de l'invention, que la molécule en question est glycosylée aux 3 sites naturels de glycosylation.

Le ou les glycanne(s) de type mannosidique ont la structure GlcNAc<sub>2</sub>-Man<sub>1</sub>. Le ou les glycanne(s) polymannosidique(s) ont la structure GlcNAc<sub>2</sub>-Man<sub>2-9</sub>, par exemple GlcNAc<sub>2</sub>-Man<sub>9</sub>, GlcNAc<sub>2</sub>-Man<sub>8</sub>, GlcNAc<sub>2</sub>-Man<sub>6</sub> ou GlcNAc<sub>2</sub>-Man<sub>5</sub>, ou GlcNAc<sub>2</sub>-Man<sub>3</sub>.

Le ou les glycanne(s) de type complexe sont biantennés et ont une structure de base GlcNAc<sub>2</sub>-Man<sub>3</sub> à laquelle sont éventuellement associés des résidus de xylose (Xyl), fucose (Fuc) et éventuellement galactose

(Gal) ou N-acétylglucosamine (GlcNAc). Ils sont normalement exempts de résidus d'acide sialique, ce composé n'ayant pas jusqu'alors été mis en évidence dans les cellules végétales.

De préférence, le glycanne complexe est de type "phytohémagglutinine" (PHA) constitué d'une structure "core"  $\text{GlcNAc}_2\text{Man}_3$  dont le mannose lié en  $\beta$  porte un résidu de  $\beta$  1,2-xylose et dont le GlcNAc proximal porte un résidu d' $\alpha$  1,3 fucose ( $\text{GlcNAc}_2$  (Fuc)  $\text{Man}_3$  (Xyl)). Ce genre de structure est fréquent pour les glycoprotéines de localisation vacuolaire ou extracellulaire. Le résidu de fucose  $\alpha$  1,3-lié peut éventuellement être absent.

Le glycanne complexe peut aussi être du type "Laccase" (Driouich et al, Regard sur la Biochimie, 1993, n°3, pp33-42) composé d'une structure de base  $\text{GlcNAc}_2$  (Fuc)  $\text{Man}_3$  (Xyl) associée à deux chaînes latérales. Chaque chaîne latérale est constituée d'un résidu de  $\beta$  1,2 GlcNAc auquel est lié un résidu d' $\alpha$  1,6 fucose et un résidu de  $\beta$  1,4 galactose.

La glycoprotéine de l'invention peut être glycosylée aux trois sites naturels (Asn-46, Asn-83 et Asn-247) ou seulement à l'un ou deux d'entre eux. Parmi les variantes préférées, l'on peut citer des glycoprotéines portant exclusivement des glycannes polymannosidiques. Ces glycoprotéines présentent une très faible immunogénicité chez l'animal parce que ce type de glycosylation existe sous forme identique chez la plante et chez l'animal.

On peut aussi citer les glycoprotéines portant exclusivement des glycannes de type complexe, par exemple deux ou trois glycannes de type PHA :  $\text{GlcNAc}_2$ (Fuc) $\text{Man}_3$ (Xyl).

La partie glucidique représente, normalement, entre 2 et 30%, par exemple 5 à 20% de la masse totale de la glycoprotéine de l'invention.

La partie protéique de la glycoprotéine correspond à la partie protéique de toute protéine ou fragment de protéine présentant une activité  $\alpha_1$ -AT. Il s'agit normalement du polypeptide mature d' $\alpha_1$ -AT humaine, mais peut également être celui d'animaux tels que le lapin ou le singe.

La séquence en acides aminés complète de l' $\alpha_1$ -AT humaine, y compris le peptide signal N-terminal (Type « M1 »), est décrite par Long et al. (Biochemistry, 1984, vol. 23, pages 4828-4837). En plus du type M1, il existe également les types M2 ou M3 (Val 213 - Asp 376 - Arg 101 ; Val 213 - Asp 376 - His 101, respectivement), ou une variante de M1, M2 ou M3 dans laquelle Val 213 est remplacé par Ala. Les protéines de l'invention peuvent comporter les séquences en acide aminés des types M1, M2 ou M3.

La séquence en acides aminés peut aussi être celle des variantes Z ou S, associées à certains troubles pulmonaires. Ces variantes se distinguent du type M1 décrit par Long et al par Glu-342 → Lys (variante Z) et Glu-264 → Val (variante S).

De manière générale, la protéine de l'invention peut comporter toute séquence en acides aminés se différenciant de la séquence naturelle par une ou plusieurs substitution(s), délétion(s) ou insertion(s) d'acides aminés. Normalement, ces variantes présentent une homologie, ou une identité, d'au moins 90% par exemple 95%, avec la séquence décrite par Long et al. (Biochemistry, 1984, vol. 23, pages 4828-4837 - Figure 1).

Parmi les variantes préférées, l'on peut citer des séquences d' $\alpha_1$ -AT dont la Met<sub>358</sub> du site actif, a été modifié, par exemple remplacé par un acide aminé différent, tel que Arg, Leu, Ala, Ile ou Val. De plus, Asp<sub>2</sub> et Asp<sub>6</sub> peuvent être remplacés par Asn.

Les variantes de la protéine de l'invention peuvent aussi comporter des fragments de la séquence décrite par Long et al. (Biochemistry, 1984, vol. 23, pages 4828-4837) et de ses analogues définis ci-dessus à condition que l'activité d'inhibiteur de protéase à sérine, et particulièrement celle de l' $\alpha_1$ -AT, soit conservée. Ces fragments sont constitués d'environ 200 à 393 acides aminés consécutifs de la séquence d' $\alpha_1$ -AT, de préférence entre 250 et 390 acides aminés. Parmi les fragments préférés, on peut citer l' $\alpha_1$ -AT mature humaine dont l'extrémité N terminale a été tronquée par la délétion d'environ 1 à 10 acides aminés.

L'extrémité C terminale peut aussi être tronquée d'environ 1 à 20 acides aminés. Le site actif de la protéine se trouvant autour des acides aminés Met-358 à Glu-363, il est important de ne pas effectuer de délétion C terminale de longueur trop importante afin de ne pas affecter le site actif.

Outre la séquence en acides aminés responsable de l'activité  $\alpha_1$ -AT proprement dite, la protéine de l'invention peut comprendre différents signaux peptidiques responsables de l'adressage de la protéine vers les différents compartiments cellulaires où sont effectuées les modifications co- et post-traductionnelles. Ces signaux sont le peptide signal N-terminal, également connu sous le nom de prépeptide, le signal de type KDEL responsable de la rétention de la protéine dans le réticulum endoplasmique et le signal d'adressage vacuolaire, ou propeptide.

En effet, dans la cellule végétale, l'adressage des protéines est basé sur le même principe que dans les cellules animales. A partir de l'ADN chromosomique, le gène est transcrit en ARN messenger, puis traduit en protéine au niveau des ribosomes. Si la protéine naissante possède un peptide signal N-terminal ou prépeptide, elle pénètre dans le réticulum

endoplasmique où ont lieu un certain nombre de maturations post-traductionnelles, notamment le clivage du peptide signal, les N-glycosylations conduisant aux glycanes polymannosidiques, et la formation des ponts disulfures. La présence d'un peptide KDEL, HEKDEL ou SEKDEL à l'extrémité C terminale provoque la rétention de la protéine dans le réticulum endoplasmique et, dans certains cas, l'accumulation de protéines dans des vésicules du réticulum endoplasmique. Lors de la rétention d'une protéine dans le réticulum endoplasmique, le signal KDEL (HDEL ou SEKDEL) n'est pas clivé et il subsiste sur la protéine mature.

En l'absence d'un signal de rétention du type KDEL, la protéine est transportée vers l'appareil de Golgi et subit, au cours de ce transport, une première modification de ses chaînes de glycanes (suppression des glucoses terminaux). Dans l'appareil de Golgi, la maturation des glycanes se poursuit par la suppression des résidus et l'addition de résidus xylose et fucose pour former les glycanes complexes. En l'absence d'un signal d'adressage vacuolaire ou "propeptide", la protéine est alors sécrétée et s'accumule dans l'espace intercellulaire. Lorsque la protéine possède à l'extrémité N-terminale ou C-terminale un propeptide, la proprotéine est dirigée vers la vacuole. A l'entrée dans la vacuole, le propeptide est clivé et certaines maturations finales de glycanes sont réalisées.

Parmi ces différents signaux, le prépeptide, responsable de l'adressage de la protéine dans le réticulum endoplasmique, est toujours présent. Il s'agit normalement d'un peptide signal N-terminal hydrophobe ayant entre 10 et 40 acides aminés. Il est normalement d'origine animale ou végétale. Il peut en effet s'agir du prépeptide naturellement associé à



l' $\alpha_1$ -AT humaine (PA), ou d'un prépeptide provenant d'une autre protéine humaine, par exemple celui de la sérum albumine humaine. Par exemple, il s'agit d'un prépeptide d'origine végétale, par exemple celui de la sporamine (PS), de la lectine d'orge, de l'extensine végétale (pEXT), de l' $\alpha$ -mating factor, des protéines végétales impliquées dans la défense contre les micro-organismes (PR1a et PRS, "pathogenesis related proteins").

Normalement, le peptide signal est clivé par un signal peptidase dès l'introduction co-traductionnelle du polypeptide naissant dans la lumière du réticulum endoplasmique (RER). La protéine mature ne comporte plus cette extension N-terminale.

La protéine de l'invention peut, outre le prépeptide, aussi comporter un signal de rétention endoplasmique, consistant en les peptides KDEL, SEKDEL ou HDEL. Ces signaux se trouvent normalement à l'extrémité C-terminale de la protéine et subsistent sur la protéine mature. La présence de ce signal sur les protéines de l'invention est avantageuse pour plusieurs raisons : d'une part, la rétention de la protéine dans le réticulum endoplasmique a tendance à augmenter les rendements en protéines recombinantes. D'autre part, la maturation de la glycosylation polymannosique en glycannes complexes n'aura pas lieu ; la protéine conserve donc la glycosylation polymannosique, minimisant le risque de réactions immunologiques indésirables lorsque la protéine sera administrée à l'homme comme médicament. Selon cette variante de l'invention, le glycoprotéine comporte donc la séquence en acides aminés d' $\alpha_1$ -AT, un signal du type KDEL, au moins un glycanne de type polymannosique, et est exempt de glycannes de type complexe. Selon l'invention, ce type de protéine peut aussi être obtenu en utilisant des mutants de plantes

incapables de fabriquer la N-acétyl glucosaminyl transférase (von Schaewen et al., Plant Physiol. 1993 102: 1109-1118), et donc incapables de produire des glycannes complexes.

La protéine de l'invention peut, outre le prépeptide, comporter aussi un signal d'adressage vacuolaire ou "propeptide". En présence d'un tel signal, la protéine est adressée aux vacuoles des tissus aqueux, par exemple les feuilles, ainsi qu'aux corps protéiques des tissus de réserve, par exemple les graines, tubercules et racines. L'adressage de la protéine vers les corps protéiques de la graine est particulièrement intéressant en raison de la capacité de la graine à accumuler des protéines, jusqu'à 40% des protéines par rapport à la matière sèche, dans des organites cellulaires dérivés des vacuoles, appelés corps protéiques et en raison de la possibilité de stocker plusieurs années les graines contenant les protéines recombinantes à l'état déshydraté.

Comme propeptide, on peut utiliser un signal d'origine animale ou végétale, les signaux végétaux étant préférés, par exemple la pro-sporamine, ou la lectine d'orge. Le propeptide peut être N-terminal ("N-terminal targeting peptide" ou NTTP), ou C terminal (CTTP) ou peut consister en une séquence interne à la protéine. Dans la mesure où le propeptide est normalement clivé dès l'entrée de la protéine dans la vacuole, il n'est pas présent dans la protéine mature. Pour un adressage vacuolaire, on peut utiliser une combinaison peptide signal et propeptide N-terminal de sporamine de patate douce (PPS).

Selon une autre variante de l'invention, la protéine de l'invention peut être sous forme de protéine de fusion, en particulier avec une protéine végétale. La protéine végétale peut être choisie afin d'augmenter le rendement ou de faciliter la sécrétion

et la purification. La protéine de fusion peut aussi être choisie afin de permettre le ciblage de l' $\alpha_1$ -AT ou de faciliter son transport.

L'objectif principal de l'invention étant de produire l' $\alpha_1$ -AT recombinante dans des cellules végétales monocotylédones, l'invention vise également les acides nucléiques codant pour les protéines, et pour les signaux d'adressage, éventuellement en association avec des séquences régulatrices adéquates, permettant l'expression de l' $\alpha_1$ -AT dans la cellule végétale.

La séquence d'acide nucléique codant pour l' $\alpha_1$ -AT, ou pour ses variantes ou les fragments définis ci-dessus, peut être de l'ADNc ou de l'ADN génomique avec ses introns. Normalement, il s'agit d'ADNc, une séquence appropriée est illustrée la Figure 1 (Long et al, Biochemistry, 1984, vol. 23, pages 4828-4837). Toute séquence dégénérée peut être utilisée. Lorsqu'il s'agit d'ADNc, des introns, provenant de préférence d'un gène végétal, peuvent être introduits artificiellement afin d'augmenter l'efficacité de l'expression de la séquence hétérologue. En effet, il a été démontré, particulièrement chez les monocotylédones que l'insertion d'un intron dans la partie 5' non traduite d'un gène, c'est-à-dire entre le site d'initiation de transcription et le site d'initiation de traduction, conduit à une amélioration de la stabilité du messager, et par conséquence, à une meilleure expression. Le ou les introns utilisé(s) de cette manière proviennent de préférence d'une monocotylédone telle que le maïs. Il s'agit de préférence, mais pas obligatoirement, du premier intron du gène.

L'invention vise également des gènes chimériques comprenant d'une part une séquence codant pour une protéine ayant une activité d' $\alpha_1$ -antitrypsine et

d'autre part, des séquences régulatrices de transcription reconnues par une cellule végétale monocotylédone. Les séquences régulatrices comprennent un ou plusieurs promoteurs d'origine végétale plus particulièrement d'origine monocotylédone, ou virale ou provenant d'*Agrobacterium tumefaciens*. Il peut s'agir d'un promoteur constitutif, par exemple le promoteur actine de riz, ou des promoteurs spécifiques de certains tissus comme le grain, ou spécifiques de certaines phases de développement de la plante monocotylédone.

Comme promoteurs spécifiques de graines, on peut citer le promoteur du gène gamma (g) zéine de maïs (Reina et al, 1990).

Il peut être envisagé d'utiliser des "enhancers" pour améliorer l'efficacité d'expression.

Les séquences régulatrices de terminaison sont d'origine végétale ou virale ou provenant d'*Agrobacterium tumefaciens*, et sont fonctionnelles dans des cellules végétales monocotylédones.

Selon une variante, le gène chimérique de l'invention comprend en outre une séquence codant pour un peptide signal permettant la sécrétion de la protéine, et éventuellement une séquence codant pour un signal de rétention endoplasmique ou pour un signal d'adressage vacuolaire.

Les promoteurs semences spécifiques sont particulièrement avantageux en association avec un signal d'adressage vacuolaire (prépropeptide).

Dans le cas où des introns hétérologues sont utilisés en combinaison avec l'ADNc d' $\alpha_1$ -AT, bien entendu, le gène chimérique comprend également ces introns.

L'invention vise également des plasmides ou vecteurs caractérisés en ce qu'ils contiennent au moins l'une des séquences d'acide nucléique, de

préférence un gène chimérique, selon l'invention. De préférence, les plasmides et vecteurs permettent la transformation stable de la plante, par exemple les plasmides Ti d'Agrobacterium, des vecteurs viraux tels que les Geminivirus.

Tous les moyens connus pour introduire de l'ADN étranger dans les cellules végétales peuvent être utilisés, par exemple Agrobacterium, électroporation, transformation de protoplastes, bombardement avec canon à particules, ou pénétration d'ADN dans des cellules comme le pollen, la microspore, la graine et l'embryon immature, suspensions cellulaires embryogènes et non embryogènes, inflorescence immature et vecteurs viraux tels que les Geminivirus. La séquence de l'invention est introduite dans un vecteur approprié avec toutes les séquences régulatrices nécessaires tels que promoteurs, terminateurs, etc... ainsi que toute séquence nécessaire pour sélectionner les transformants.

L'intégration de l'acide nucléique dans le génome peut éventuellement avoir lieu par recombinaison homologue. Dans ce cas, l'acide nucléique transformant comporte, sur chaque extrémité, une séquence homologue aux séquences qui jouxtent le site d'insertion souhaité dans le génome.

L'invention concerne également les cellules végétales monocotylédones transformées avec les séquences de l'invention, et capable de produire une ou plusieurs protéine(s) ayant une activité d'inhibiteur de protéase à serine et, en particulier, de l' $\alpha_1$ -AT.

Il peut s'agir de cultures de cellules végétales in vitro, par exemple en milieu liquide. Différents modes de culture ("batch", "fed batch" ou en continu) pour ce type de cellules sont actuellement à l'étude. Les cultures en "batch" sont comparables à celles

effectuées en erlenmeyer dans la mesure où le milieu n'est pas renouvelé, les cellules ne disposent ainsi que d'une quantité limitée d'éléments nutritifs. La culture en "fed batch" correspond quant à elle à une culture en "batch" avec une alimentation programmée en substrat. Pour une culture en continu, les cellules sont alimentées en permanence avec du milieu nutritif. Un volume égal du mélange biomasse-milieu est ôté afin de maintenir le volume du réacteur constant. Les quantités de biomasse végétale envisageables avec des cultures en bioréacteurs sont variables selon l'espèce végétale, le mode de culture et le type de bioréacteur.

Les cellules de l'invention peuvent aussi être immobilisées, ce qui permet d'obtenir une production constante et prolongée d' $\alpha_1$ -AT. La séparation de l' $\alpha_1$ -AT et la biomasse végétale est aussi facilitée. Comme méthode d'immobilisation, on peut citer l'immobilisation en billes d'alginate, d'agar, à l'intérieur de mousse de polyuréthane, ou bien encore dans des fibres creuses.

Au lieu de produire l' $\alpha_1$ -AT de l'invention par culture de cellules végétales, on peut régénérer des plantes chimériques ou transgéniques à partir d'explants transformés, en ayant recours à des techniques connues en soi.

Comme plantes préférées, on peut citer toutes les plantes monocotylédones. On peut citer les céréales telles que le blé, le maïs, le riz, l'orge, le sorgho et l'avoine, mais aussi la canne à sucre, l'asperge et la banane.

L'invention concerne également les graines des plantes transgéniques capables de produire l' $\alpha_1$ -AT ainsi que leurs descendances.

Selon le procédé de l'invention, les protéines sont récupérées, et éventuellement purifiées, afin de

permettre leur utilisation dans un grand nombre d'applications.

Les méthodes de récupération (extraction) et de purification des protéines sont choisies en fonction de la méthode de production, c'est-à-dire cellules en culture ou plantes entières, et éventuellement de l'adressage mis en oeuvre.

L'extraction est normalement faite par broyage des tissus, par exemple feuilles ou grains, dans un tampon approprié, filtrage du broyat, précipitation des protéines dans le surnageant, centrifugation et reprise du culot dans un tampon approprié avec dialyse. Une étape de purification partielle peut aussi être effectuée à ce stade par chromatographie sur colonne d'échangeuse d'ions.

L'invention vise aussi des anticorps monoclonaux ou polyclonaux capables de reconnaître spécifiquement la protéine de l'invention. Ces anticorps peuvent être utilisés dans la purification des protéines de l'invention et dans le suivi des patients.

L'invention vise en outre un produit pharmaceutique comprenant une ou plusieurs protéine(s) de l'invention en association avec un excipient acceptable du point de vue physiologique.

La composition pharmaceutique de l'invention peut être administrée sous forme d'aérosol (particulièrement appropriée, s'il s'agit de l' $\alpha_1$ -AT non-glycosylée) pour injection nasale, ou par voie intraveineuse, sous-cutanée, topique ou percutanée.

Il est également possible d'utiliser un implant sous-cutané des cellules végétales transformées sans purification de l' $\alpha_1$ -AT, isolées dans une membrane semi-perméable laissant passer la protéine active. La ou les protéine(s) de l'invention peut ou peuvent être administrée(s) en forme de liposome(s), encapsidée(s) ou conjuguée(s) à une molécule de ciblage.

L'invention vise également l'utilisation de la protéine de l'invention pour la préparation d'un médicament pour le traitement de conditions liées à une déficience en  $\alpha_1$ -antitrypsine. Les maladies qui sont susceptibles d'être traitées comprennent la mucoviscidose, le choc septique, et l'emphysème pulmonaire. La composition pharmaceutique de l'invention peut aussi être utilisée comme anti-rhumatoïde grâce à son effet anti-collagénase.

Les protéines de l'invention peuvent également être utilisées en cosmétologie. Leur activité anti-collagénase empêche la dégradation du tissu conjonctif et en particulier, du collagène et de l'élastine. Selon cette variante de l'invention, les protéines peuvent être purifiées ou peuvent être utilisées sous forme d'extraits de toute ou partie de la plante ou de la graine ou d'extraits de culture cellulaire.

L'invention vise l'utilisation de la protéine de l'invention dans une application industrielle.

Les protéines de l'invention peuvent aussi être utilisées comme réactifs anti-protéases sous forme de préparations enzymatiques industrielles.

L'invention concerne également un extrait de plante contenant 1 à 90% environ d'une protéine selon l'invention, et notamment l' $\alpha_1$ -antitrypsine. Cet extrait peut être utilisé comme médicament, notamment en thérapie de conditions liées à une déficience en  $\alpha_1$ -AT, notamment l'emphysème pulmonaire, la mucoviscidose, le choc septique ou comme anti-rhumatoïde. L'invention concerne également une composition pharmaceutique comprenant un tel extrait de plante en association avec un excipient acceptable du point de vue physiologique, de préférence pour l'administration par voie orale ou nasale, par exemple sous forme d'aérosol.



Des modes de réalisation préférés de l'invention sont présentés dans les revendications, en particulier :

- 1°) Un procédé de production d' $\alpha_1$ -antitrypsine ( $\alpha_1$ -AT), ou de variantes de celle-ci caractérisé par :
  - i) l'introduction dans des cellules végétales monocotylédones, d'un acide nucléique comportant une séquence codant pour l' $\alpha_1$ -antitrypsine ( $\alpha_1$ -AT) ou pour une variante de celle-ci, les variantes se différenciant de l' $\alpha_1$ -AT humaine naturelle par une ou plusieurs substitution(s), délétion(s) ou insertion(s) d'acides aminés, et éventuellement une séquence codant pour un agent sélectif ;
  - ii) sélection des cellules ayant intégré de manière stable l'acide nucléique ;
  - iii) propagation des cellules transformées en culture, ou régénération de plantes entières chimériques ou transgéniques ;
  - iv) récupération et éventuellement purification de l' $\alpha_1$ -antitrypsine ou de ses variantes, ainsi produites.
- 2°) Un procédé selon le procédé énoncé au 1°) ci-dessus caractérisé en ce que l' $\alpha_1$ -antitrypsine ou ses variantes présentent une activité d'inhibition de protéases à sérine, de préférence une activité de l' $\alpha_1$ -AT humaine.
- 3°) Un procédé selon le procédé énoncé au 2°) ci-dessus caractérisé en ce que l'acide nucléique introduit dans les cellules végétales code pour la protéine illustrée dans la figure 1, ou pour une protéine présentant au moins 90% d'homologie avec celle-ci, ou encore un fragment de l'une de ces protéines ayant une longueur comprise entre 200 et 393 acides aminés.
- 4°) Un procédé selon le procédé énoncé au 3°) ci-dessus caractérisé en ce que l'acide nucléique comporte en outre une séquence codant pour un

peptide signal ou "prépeptide", et éventuellement une séquence codant pour un signal de rétention endoplasmique ou pour un signal d'adressage vacuolaire.

- 5°) Un procédé selon l'un quelconque des procédés énoncés au 3°) ou 4°) ci-dessus caractérisé en ce que l'acide nucléique comporte en outre des séquences régulatrices de transcription reconnues par une cellule végétale monocotylédone.
- 6°) Une protéine présentant une activité d'inhibiteur de protéase à sérine, de préférence une activité de l' $\alpha_1$ -antitrypsine humaine, caractérisée en ce qu'elle est susceptible d'être produite par le procédé selon l'un quelconque des procédés énoncés aux 1°) à 5°) ci-dessus.
- 7°) Une protéine selon la protéine énoncée au 6°) ci-dessus caractérisée en ce qu'elle est N-glycosylée par au moins un glycanne de type mannosidique ou polymannosidique, et/ou par au moins un glycanne de type complexe comportant éventuellement au sein de sa structure un ou plusieurs résidus de xylose et éventuellement un ou plusieurs résidus de fucose.
- 8°) Une protéine selon la protéine énoncée au 7°) ci-dessus comportant la séquence en acides aminés illustrée dans la figure 1, ou une séquence présentant au moins 90% d'homologie avec cette séquence, ou encore un fragment de l'une de ces séquences ayant une longueur comprise entre 200 et 393 acides aminés.
- 9°) Une protéine selon la protéine énoncée au 8°) ci-dessus caractérisée en ce que le glycanne est lié N-glycosidiquement à Asn-46, Asn-83 ou Asn-247.
- 10°) Une protéine selon l'une quelconque des protéines énoncées précédemment caractérisée en ce qu'elle porte à la fois un ou plusieurs glycannes de type

polymannosique et un ou plusieurs glycanes de type complexe.

- 11°) Une protéine selon l'une quelconque des protéines énoncées précédemment caractérisée en ce que le glycanne polymannosique présente la structure GlcNAc2-Man5-9, par exemple GlcNAc2-Man9.
- 12°) Une protéine selon l'une quelconque des protéines énoncées précédemment caractérisée en ce que le glycanne complexe présente la structure GlcNAc2 (Fuc) Man3 (Xyl).
- 13°) Une protéine selon l'une des protéines énoncées aux 6°) à 12°) ci-dessus caractérisée en ce que la séquence en acides aminés comporte en outre un peptide signal N-terminal, permettant une sécrétion de la protéine et éventuellement un signal responsable de la rétention de la protéine dans le réticulum endoplasmique, ou un signal reconnu par une cellule végétale comme étant un signal d'adressage vacuolaire.
- 14°) Une protéine selon la protéine énoncée au 13°) ci-dessus caractérisée en ce que le peptide signal est le prépeptide de la sporamine, de la lectine d'orge, de l'extensine végétale, de l' $\alpha$ -mating factor, des protéines de pathogénèse.
- 15°) Une protéine selon la protéine énoncée au 13°) ci-dessus caractérisée en ce que le peptide signal est celui de l' $\alpha_1$ -antitrypsine native.
- 16°) Une protéine selon la protéine énoncée au 13°) ci-dessus caractérisée en ce que le signal responsable de la rétention de la protéine dans le réticulum endoplasmique est choisi parmi les peptides KDEL, SEKDEL et HDEL.
- 17°) Une protéine selon la protéine énoncée au 13°) ci-dessus caractérisée en ce que le signal d'adressage vacuolaire est d'origine végétale, tel

que celui de la sporamine, ou celui de la lectine d'orge.

- 18°) Une protéine selon la protéine énoncée au 6°) ci-dessus caractérisée en ce qu'elle est exempte de chaînes N-glycosidiques.
- 19°) Un acide nucléique comportant : i) une séquence codant pour l' $\alpha_1$ -AT ou pour une variante de celle-ci, et ii) une séquence d'adressage choisie parmi une séquence codant pour un peptide signal "pré" d'origine végétale et éventuellement une séquence codant pour un signal de rétention ou pour un signal d'adressage vacuolaire, et iii) des séquences régulatrices de transcription reconnues par une cellule végétale monocotylédone.
- 20°) Un acide nucléique selon l'acide nucléique énoncé au 19°) ci-dessus caractérisé en ce que les séquences régulatrices comprennent un ou plusieurs promoteur(s) d'origine végétale ou virale.
- 21°) Un acide nucléique selon l'un quelconque des acides nucléiques énoncés aux 19°) et 20°) ci-dessus caractérisé en ce que la séquence codant pour l' $\alpha_1$ -AT est un ADNc codant pour l' $\alpha_1$ -AT mature.
- 22°) Un vecteur comprenant un acide nucléique selon l'un quelconque des acides nucléiques énoncés aux 19°) à 21°).
- 23°) Des cellules végétales monocotylédones transformées de manière stable par un acide nucléique selon l'un quelconque des acides nucléiques énoncés aux 19°) à 21°) ci-dessus, et capable de produire une ou plusieurs protéine(s) ayant une activité d'inhibiteur de protéase à sérine, de préférence une activité d' $\alpha_1$ -antitrypsine.
- 24°) Des cellules végétales monocotylédones selon les cellules énoncées au 23°) ci-dessus caractérisées

en ce qu'il s'agit d'une culture de cellules végétales, par exemple en milieu liquide ou immobilisées, ou une culture de racines.

- 25°) Des cellules végétales selon les cellules énoncées au 24°) ci-dessus caractérisées en ce qu'il s'agit de cellules faisant partie d'une plante monocotylédone transformée.
- 26°) Une plante monocotylédone chimérique ou transgénique capable de produire une protéine ayant une activité d'inhibiteur de protéase à sérine, de préférence une activité d' $\alpha_1$ -antitrypsine, caractérisée en ce qu'elle comporte des cellules selon les cellules énoncées au 23°) ci-dessus.
- 27°) Des graines de plante transgénique selon la plante énoncée au 26°) ci-dessus.
- 28°) Des anticorps monoclonaux ou polyclonaux capables de reconnaître spécifiquement la protéine selon l'une des protéines énoncées aux 6°) à 18°) ci-dessus.
- 29°) un produit pharmaceutique comprenant une ou plusieurs protéine(s) selon l'une quelconque des protéines énoncées aux 6°) à 18°) ci-dessus en association avec un excipient acceptable du point de vue physiologique.
- 30°) Une protéine ayant une activité d' $\alpha_1$ -AT selon l'une quelconque des protéines énoncées aux 6°) à 18°) ci-dessus pour l'utilisation comme médicament.
- 31°) Une protéine selon la protéine énoncée au 30°) ci-dessus pour une utilisation en thérapie de conditions liées à une déficience en  $\alpha_1$ -AT, notamment l'emphysème pulmonaire, la mucoviscidose, le choc septique, ou comme anti-rhumatoïde.

- 32°) Une utilisation d'une protéine selon l'une quelconque des protéines énoncées aux 6°) à 18°) ci-dessus pour la préparation d'un médicament pour le traitement de conditions liées à une déficience en  $\alpha_1$ -AT.
- 33°) Une utilisation d'une protéine selon l'une quelconque des protéines énoncées aux 6°) à 18°) ci-dessus dans la préparation de produits cosmétiques, ou de réactifs chimiques.
- 34°) Un extrait de plante contenant 1 à 90% environ d'une protéine selon l'une quelconque des protéines énoncées aux 6°) à 18°) ci-dessus, et notamment l' $\alpha_1$ -antitrypsine.
- 35°) Un extrait de plante selon l'extrait énoncé au 34°) ci-dessus pour l'utilisation comme médicament.
- 36°) Un extrait de plante selon l'extrait énoncé au 35°) ci-dessus pour une utilisation en thérapie de conditions liées à une déficience en  $\alpha_1$ -AT, notamment l'emphysème pulmonaire, la mucoviscidose, le choc septique ou comme anti-rhumatoïde.
- 37°) Une composition pharmaceutique comprenant un extrait de plante selon l'extrait énoncé au 34°) ci-dessus en association avec un excipient acceptable du point de vue physiologique, de préférence pour l'administration par voie orale ou nasale, par exemple sous forme d'aérosol.

La Figure 1 illustre l'ADNc et la séquence en acide aminés de l' $\alpha_1$ -AT humaine (type M). L'extrémité  $\text{NH}_2$  de la protéine mature (Glu) est indiqué comme +1. Les acides aminés -1 à -24 représentent le peptide signal N-terminal de la protéine immature. Le site réactif (Met) est indiqué par un cercle solide (•). Quatre sites de glycosylation potentiels sont présents dans la protéine, indiqués par des diamants noirs (◆), mais

seulement trois d'entre eux portent des chaînes glycosidiques (Long et al Supra).

La Figure 2 illustre la carte schématique du plasmide pUC-AAT. L'ADNc de AAT est inséré au site PstI de pUC18. PA (boîte en pointillés) correspond à la séquence du peptide signal AAT. AAT (boîte en noir) correspond à la séquence codant l' $\alpha$ 1-antitrypsine humaine. Ap (boîte en blanc) correspond au gène conférant la résistance à l'ampicilline.

#### EXEMPLES

##### I. CONSTRUCTIONS

L' $\alpha$ 1-antitrypsine humaine (AAT), une glycoprotéine de 53 kDa, est naturellement synthétisée sous forme d'un précurseur de 418 acides aminés. La protéine mature AAT est constituée de 394 acides aminés et est clivée au site Ala - Glu de son peptide signal de 24 acides aminés. La séquence complète du cDNA ainsi que la séquence peptidique sont décrites par Long et al. (Biochemistry, 1984, vol. 23, pages 4828-4837).

La séquence complète de l'ADNc est renfermée dans le plasmide pUC18 commercialisé par Clontech. Ce plasmide est appelé pUC-AAT (Figure 2).

L'ADNc de AAT, fragment d'ADN digéré par StuI et SalI et isolé à partir de pUC18, a été cloné aux sites XbaI traité à la Klenow et SalI du vecteur pBSIISK+ commercialisé par Stratagene Cloning Systems. Le plasmide résultant a été appelé pBIOC30.

**A. CONSTRUCTION DES VECTEURS pBIOC31 ET pBIOC32.  
MODIFICATION DE L'EXTREMITÉ 3' DU cDNA CODANT POUR  
L' $\alpha$ 1-ANTITRYPSINE**

**A.1. CONSTRUCTION DU PLASMIDE pBIOC31 NE  
CONTENANT QUE LE PREMIER CODON STOP DE L' $\alpha$ 1-  
ANTITRYPSINE**

Pour favoriser l'expression dans les plantes, seul le premier codon stop (TAA) de la protéine  $\alpha$ 1-antitrypsine humaine (AAT) a été préservé. Le vecteur pBIOC30 a donc été modifié à l'extrémité 3' du cDNA codant pour AAT.

Le fragment EcoRV-SalI de pBIOC30 a été remplacé par le fragment EcoRV-SalI issu de l'amplification PCR réalisée sur pBIOC30 à l'aide des 2 oligodésoxynucléotides, 5' CCACGATATCATCACCAAGTTCC 3' (contenant le site unique EcoRV) et 5' cggtcgacgaattcCAGTTATTTTGGGTGGGATTC 3' (contenant les sites SalI et EcoRI, et le premier codon stop de l'AAT). L'amplification PCR du fragment EcoRV-SalI a été réalisée dans 100  $\mu$ l de milieu réactionnel comprenant 10  $\mu$ l de tampon Taq DNA polymérase x10 (500 mM KCl, 100 mM Tris-HCl, pH9,0 et 1% Triton x100), 6  $\mu$ l de 25 mM MgCl<sub>2</sub>, 3  $\mu$ l de 10 mM dNTP (dATP, dCTP, dGTP et dTTP), 100 pM de chacun des 2 oligodésoxynucléotides décrits ci-dessus, 5 ng d'ADN matrice (vecteur pBIOC30), 2,5 U de Taq DNA polymérase (Promega) et 2 gouttes d'huile de vaseline. L'ADN a été dénaturé à 94°C pendant 5 min., soumis à 30 cycles constitués chacun de 1 min. de dénaturation à 94°C, 1 min. d'hybridation à 70°C et de 1 min. d'élongation à 72°C, puis l'élongation à 72°C a été poursuivie pendant 5 min. Cette réaction PCR a été réalisée dans la machine "DNA Thermal Cyclor" de PERKIN ELMER CETUS. L'huile a été éliminée par extraction au chloroforme.



Puis, les fragments d'ADN du milieu réactionnel ont été précipités en présence de 1/10 de volume de 3M acétate de sodium pH4,8 et de 2,5 volumes d'éthanol absolu à -80°C pendant 30 min., centrifugés à 12000g pendant 30 min., lavés à l'éthanol 70%, séchés et digérés par les 2 enzymes de restriction EcoRV et SalI. Les fragments d'ADN digérés issus de PCR ont été purifiés par électrophorèse sur gel d'agarose 2%, électroélués (Sambrook et al., 1989), précipités en présence de 1/10 de volume de 3M acétate de sodium pH4,8 et de 2,5 volumes d'éthanol absolu à -80°C pendant 30 min., centrifugés à 12000g pendant 30 min., lavés à l'éthanol 70%, séchés, puis ligués à l'ADN plasmidique pBIOC30 digéré par EcoRV et SalI, purifié par électrophorèse sur gel d'agarose 0,8%, électroélué, soumis à la précipitation alcoolique, séché et déphosphorylé par l'enzyme phosphatase alcaline d'intestin de veau (Boehringer Mannheim) selon les recommandations du fabricant. La ligation a été réalisée avec 100 ng du vecteur déphosphorylé décrit ci-dessus et 50 ng de fragments d'ADN digérés issus de l'amplification PCR décrits ci-dessus dans un milieu réactionnel de 10 µl en présence de 1 µl de tampon T4 DNA ligase x 10 (Amersham) et de 2,5 U d'enzyme T4 DNA ligase (Amersham) à 14°C pendant 16 heures. Les bactéries, *Escherichia coli* DH5α rendues préalablement compétentes, ont été transformées (Hanahan, 1983). L'ADN plasmidique des clones obtenus, sélectionnés sur 50 µg/ml ampicilline, a été extrait selon la méthode de la lyse alcaline (Birnboim et Doly, 1979) et analysé par digestion enzymatique par des enzymes de restriction. L'ADN plasmidique de certains des clones retenus a été vérifié par séquençage à l'aide du kit de séquençage T7™ commercialisé par Pharmacia selon la méthode des

didésoxynucléotides (Sanger et al., 1977). Le plasmide résultant a été appelé pBIOC31.

A.2. CONSTRUCTION DU PLASMIDE pBIOC32 CONTENANT LE PREMIER CODON STOP DE L' $\alpha$ -ANTITRYPSINE PRECEDE DU SIGNAL KDEL

La séquence codant le signal KDEL (Lys-Asp-Glu-Leu) placée à l'extrémité C-terminale de la séquence codant la protéine mature  $\alpha$ 1-antitrypsine humaine en amont du codon stop combinée à la présence de la séquence codant le peptide signal N-terminal de la sporamine A des racines tubérisées de patate douce permet un adressage dans le réticulum endoplasmique.

Pour insérer la séquence codant le signal KDEL avant le premier codon stop (TAA) de la protéine mature  $\alpha$ 1-antitrypsine humaine (AAT) qui seul sera préservé, le vecteur pBIOC30 a été modifié à l'extrémité 3' de l'ADNc codant pour AAT. La présence d'un codon Lys juste avant le codon stop de AAT, s'est soldée par l'introduction de la séquence codant le signal DEL uniquement.

Le fragment EcoRV-SalI de pBIOC30 a été remplacé par le fragment EcoRV-SalI issu de l'amplification PCR réalisée sur pBIOC30 à l'aide des 2 oligodésoxynucléotides commandés chez Genset, 5' CCACGATATCATCACCAAGTTCC 3' (contenant le site unique EcoRV) et 5' cggtcgacgaattcCAGTTAtagctcatcTTTTTGGGTGGG 3' (contenant les sites SalI et EcoRI, et la séquence KDEL et le premier codon stop de AAT mature), en suivant le protocole d'amplification PCR décrit ci-dessus dans le paragraphe a.1. Après double digestion enzymatique par EcoRV et SalI, les fragments d'ADN issus de l'amplification PCR ont été purifiés par électrophorèse sur gel d'agarose 2%, électroélus (Sambrook et al., 1989), précipités en présence de

1/10 de volume de 3M acétate de sodium pH4,8 et de 2,5 volumes d'éthanol absolu à -80°C pendant 30 min., centrifugés à 12000g pendant 30 min., lavés à l'éthanol 70%, séchés, puis ligués à l'ADN plasmidique de pBIOC30 doublement digéré par EcoRV et SalI, purifié par électrophorèse sur gel d'agarose 0,8%, électroélué (Sambrook et al., 1989), soumis à la précipitation alcoolique, séché et déphosphorylé par l'enzyme phosphatase alcaline d'intestin de veau (Boehringer Mannheim) selon les recommandations du fabricant. La ligation a été réalisée avec 100 ng du vecteur déphosphorylé décrit ci-dessus et 50 ng de fragments d'ADN digérés issus de l'amplification PCR décrits ci-dessus dans un milieu réactionnel de 10 µl en présence de 1 µl de tampon T4 DNA ligase x 10 (Amersham) et de 2,5 U d'enzyme T4 DNA ligase (Amersham) à 14°C pendant 16 heures. Les bactéries, *Escherichia coli* DH5α rendues préalablement compétentes, ont été transformées (Hanahan, 1983). L'ADN plasmidique des clones obtenus, sélectionnés sur 50 µg/ml ampicilline, a été extrait selon la méthode de la lyse alcaline (Birnboim et Doly, 1979) et analysé par digestion enzymatique par des enzymes de restriction. L'ADN plasmidique de certains des clones retenus a été vérifié par séquençage à l'aide du kit de séquençage T7<sup>TM</sup> commercialisé par Pharmacia selon la méthode des didésoxynucléotides (Sanger et al., 1977). Le plasmide résultant a été appelé pBIOC32.

**B. CONSTRUCTION DES VECTEURS pBIOC33, pBIOC34, pBIOC35 ET pBIOC74. MODIFICATION DE L'EXTREMITÉ 5' DU cDNA CODANT POUR L'α1-ANTITRYPSINE**

Dans l'une des constructions, la séquence contenant l'α1-antitrypsine humaine mature précédée

d'une partie de celle de son peptide signal naturel PA (11 derniers codons des 24 nécessaires) est contenue dans pBIOC31. Le clone modifié résultant a été appelé pBIOC74.

Par ailleurs, la séquence codant pour le peptide signal naturel PA de l' $\alpha$ 1-antitrypsine humaine (24 premiers codons) a été remplacé par celle codant pour un signal d'adressage d'origine végétale en respectant les phases ouvertes de lecture des séquences codant pour le signal d'adressage apporté et la protéine mature  $\alpha$ 1-antitrypsine humaine. Ce signal d'adressage est soit un peptide signal (PS) qui permettrait une sécrétion de la protéine dans le milieu extracellulaire de l'espace intercellulaire, soit un prépropeptide c'est-à-dire un peptide signal suivi des séquences N-terminales d'adressage vacuolaire (PPS) qui adresserait la protéine dans la vacuole. Les séquences PS et PPS constituées respectivement de 23 et 37 acides aminés, sont celles d'une protéine de réserve des racines tubérisées de patate douce : la sporamine A (Murakami et al., 1986 ; Matsuoka et Nakamura, 1991). Le plasmide pBIOC31 modifié contenant la séquence codant pour l' $\alpha$ 1-antitrypsine humaine (AAT) fusionnée soit à PS (PS-AAT), soit à PPS (PPS-AAT) a été appelé pBIOC33 et pBIOC34 respectivement. Le plasmide pBIOC32 modifié contenant la séquence codant pour l' $\alpha$ 1-antitrypsine humaine (AAT) fusionnée à PS (PS-AAT-KDEL) a été appelé pBIOC35.

#### B.1. CONSTRUCTION DU PLASMIDE pBIOC74 CONTENANT PA-AAT.

Le plasmide pBIOC31 a été digéré doublement par SacI et BamHI pour supprimer les 11 derniers codons de la séquence codant pour le peptide signal naturel PA et le premier codon de l' $\alpha$ 1-antitrypsine humaine

mature. Cette séquence a été remplacée par la séquence entière codant pour le peptide signal naturel PA de 24 acides aminés (ATG CCG TCT TCT GTC TCG TGG GGC ATC CTC CTG CTG GCA GGC CTG TGC TGC CTG GTC CCT GTC TCC CTG GCT) fusionnée au premier codon de la protéine AAT mature (Glu). La séquence "PS - 3 premiers codons de la protéine AAT mature" a été amplifiée par PCR à partir du plasmide pUC18-AAT à l'aide des 2 oligodésoxynucléotides, 5' gggagctcgaattcaacaATG CCG TCT TCT GTC TCG TG 3' (contenant les sites EcoRI et SacI et le signal traductionnel de Kozack) et 5' G GGG ATC CTC AGC CAG GG3' (contenant le site BamHI de la séquence codant pour la protéine AAT mature), en suivant le protocole d'amplification PCR décrit ci-dessus dans le paragraphe a.1. Après double digestion enzymatique par SacI et BamHI, les fragments d'ADN issus de l'amplification PCR ont été purifiés par électrophorèse sur gel d'agarose 2%, électroélus (Sambrook et al., 1989), précipités en présence de 1/10 de volume de 3M acétate de sodium pH4,8 et de 2,5 volumes d'éthanol absolu à -80°C pendant 30 min., centrifugés à 12000g pendant 30 min., lavés à l'éthanol 70%, séchés, puis ligués à l'ADN plasmidique de pBIOC31 doublement digéré par SacI et BamHI, purifié par électrophorèse sur gel d'agarose 0,8%, électroélué (Sambrook et al., 1989), soumis à la précipitation alcoolique, séché et déphosphorylé par l'enzyme phosphatase alcaline d'intestin de veau (Boehringer Mannheim) selon les recommandations du fabricant. La ligation a été réalisée avec 100 ng du vecteur déphosphorylé décrit ci-dessus et 50 ng de fragments d'ADN digérés issus de l'amplification PCR décrits ci-dessus dans un milieu réactionnel de 10 µl en présence de 1 µl de tampon T4 DNA ligase x 10 (Amersham) et de 2,5 U d'enzyme T4 DNA ligase (Amersham) à 14°C pendant 16 heures. Les bactéries,

*Escherichia coli* DH5 $\alpha$  rendues préalablement compétentes, ont été transformées (Hanahan, 1983). L'ADN plasmidique des clones obtenus, sélectionnés sur 50  $\mu$ g/ml ampicilline, a été extrait selon la méthode de la lyse alcaline (Birnboim et Doly, 1979) et analysé par digestion enzymatique par des enzymes de restriction. L'ADN plasmidique de certains des clones retenus a été vérifié par séquençage à l'aide du kit de séquençage T7<sup>TM</sup> commercialisé par Pharmacia selon la méthode des didésoxynucléotides (Sanger et al., 1977). Les séquences du PA et de AAT mature ont été clonées en maintenant leurs phases de lecture ouverte. La séquence de clivage entre les séquences PA et AAT mature est Ala-Glu. Le plasmide résultant a été appelé pBIOC74.

#### B.2. CONSTRUCTION DU PLASMIDE pBIOC33 CONTENANT PS-AAT.

Le plasmide pBIOC31 a été digéré doublement par SacI et BamHI pour supprimer la séquence codant pour le peptide signal naturel de l' $\alpha$ 1-antitrypsine humaine. Cette séquence a été remplacée par celle codant pour le peptide signal PS de 23 acides aminés (ATG AAA GCC TTC ACA CTC GCT CTC TTC TTA GCT CTT TCC CTC TAT CTC CTG CCC AAT CCA GCC CAT TCC) fusionnée aux 3 premiers codons de la protéine AAT mature (Glu-Asp-Pro). La séquence "PS - 3 premiers codons de la protéine AAT mature" a été amplifiée par PCR à partir du plasmide pMAT103 (Matuoka et Nakamura, 1991) à l'aide des 2 oligodésoxynucléotides, 5' gcgagctcgaattcaacaATG AAA GCC TTC ACA CTC GC 3' (contenant les sites EcoRI et SacI et le signal traductionnel de Kozack) et 5' CT GGG ATC CTC GGA ATG GGC TGG ATT GGG CAG G 3' (contenant le site BamHI de la séquence codant pour la protéine AAT mature), en

suivant le protocole d'amplification PCR décrit ci-dessus dans le paragraphe a.1. Après double digestion enzymatique par SacI et BamHI, les fragments d'ADN issus de l'amplification PCR ont été purifiés par électrophorèse sur gel d'agarose 2%, électroélués (Sambrook et al., 1989), précipités en présence de 1/10 de volume de 3M acétate de sodium pH4,8 et de 2,5 volumes d'éthanol absolu à -80°C pendant 30 min., centrifugés à 12000g pendant 30 min., lavés à l'éthanol 70%, séchés, puis ligués à l'ADN plasmidique de pBIOC31 doublement digéré par SacI et BamHI, purifié par électrophorèse sur gel d'agarose 0,8%, électroélué (Sambrook et al., 1989), soumis à la précipitation alcoolique, séché et déphosphorylé par l'enzyme phosphatase alcaline d'intestin de veau (Boehringer Mannheim) selon les recommandations du fabricant. La ligation a été réalisée avec 100 ng du vecteur déphosphorylé décrit ci-dessus et 50 ng de fragments d'ADN digérés issus de l'amplification PCR décrits ci-dessus dans un milieu réactionnel de 10  $\mu$ l en présence de 1  $\mu$ l de tampon T4 DNA ligase x 10 (Amersham) et de 2,5 U d'enzyme T4 DNA ligase (Amersham) à 14°C pendant 16 heures. Les bactéries, *Escherichia coli* DH5 $\alpha$  rendues préalablement compétentes, ont été transformées (Hanahan, 1983). L'ADN plasmidique des clones obtenus, sélectionnés sur 50  $\mu$ g/ml ampicilline, a été extrait selon la méthode de la lyse alcaline (Birnboim et Doly, 1979) et analysé par digestion enzymatique par des enzymes de restriction. L'ADN plasmidique de certains des clones retenus a été vérifié par séquençage à l'aide du kit de séquençage T7<sup>TM</sup> commercialisé par Pharmacia selon la méthode des didésoxynucléotides (Sanger et al., 1977). Les séquences du PS et de AAT mature ont été clonées en maintenant leurs phases de lecture ouverte. La séquence de clivage entre les séquences PS et AAT

mature est Ser-Glu. Le plasmide résultant a été appelé pBIOC33.

### B.3. CONSTRUCTION DU PLASMIDE pBIOC34 CONTENANT PPS-AAT.

Le plasmide pBIOC31 a été digéré doublement par SacI et BamHI pour supprimer la séquence codant pour le peptide signal naturel de l' $\alpha$ 1-antitrypsine humaine. Cette séquence a été remplacée par celle codant pour le prépropeptide PPS N-terminal de 37 acides aminés (ATG AAA GCC TTC ACA CTC GCT CTC TTC TTA GCT CTT TCC CTC TAT CTC CTG CCC AAT CCA GCC CAT TCC AGG TTC AAT CCC ATC CGC CTC CCC ACC ACA CAC GAA CCC GCC) fusionnée aux 3 premiers codons de la protéine AAT mature (Glu-Asp-Pro). La séquence "PPS - 3 premiers codons de la protéine AAT mature" a été amplifiée par PCR à partir du plasmide pMAT103 (Matuoka et Nakamura, 1991) à l'aide des 2 oligodésoxynucléotides, 5' gcgagctcgaattcaacaATG AAA GCC TTC ACA CTC GC 3' (contenant les sites EcoRI et SacI et le signal traductionnel de Kozack) et 5' CT GGG ATC CTC GGC GGG TTC GTG TGT GGT TG 3' (contenant le site BamHI de la séquence de la protéine AAT mature), en suivant le protocole d'amplification PCR décrit ci-dessus dans le paragraphe a.1. Après double digestion enzymatique par SacI et BamHI, les fragments d'ADN issus de l'amplification PCR ont été purifiés par électrophorèse sur gel d'agarose 2%, électroélués (Sambrook et al., 1989), précipités en présence de 1/10 de volume de 3M acétate de sodium pH4,8 et de 2,5 volumes d'éthanol absolu à -80°C pendant 30 min., centrifugés à 12000g pendant 30 min., lavés à l'éthanol 70%, séchés, puis ligués à l'ADN plasmidique de pBIOC31 doublement digéré par SacI et BamHI, purifié par électrophorèse sur gel d'agarose



0,8%, électroélué (Sambrook et al., 1989), soumis à la précipitation alcoolique, séché et déphosphorylé par l'enzyme phosphatase alcaline d'intestin de veau (Boehringer Mannheim) selon les recommandations du fabricant. La ligation a été réalisée avec 100 ng du vecteur déphosphorylé décrit ci-dessus et 50 ng de fragments d'ADN digérés issus de l'amplification PCR décrits ci-dessus dans un milieu réactionnel de 10  $\mu$ l en présence de 1  $\mu$ l de tampon T4 DNA ligase x 10 (Amersham) et de 2,5 U d'enzyme T4 DNA ligase (Amersham) à 14°C pendant 16 heures. Les bactéries, *Escherichia coli* DH5 $\alpha$  rendues préalablement compétentes, ont été transformées (Hanahan, 1983). L'ADN plasmidique des clones obtenus, sélectionnés sur 50  $\mu$ g/ml ampicilline, a été extrait selon la méthode de la lyse alcaline (Birnboim et Doly, 1979) et analysé par digestion enzymatique par des enzymes de restriction. L'ADN plasmidique de certains des clones retenus a été vérifié par séquençage à l'aide du kit de séquençage T7<sup>TM</sup> commercialisé par Pharmacia selon la méthode des didésoxynucléotides (Sanger et al., 1977). Les séquences du PPS et de AAT mature ont été clonées en maintenant leurs phases de lecture ouverte. La séquence de clivage entre les séquences PPS et AAT mature est Ala-Glu. Le plasmide résultant a été appelé pBIOC34.

#### B.4. CONSTRUCTION DU PLASMIDE pBIOC35 CONTENANT PS-AAT-KDEL.

Le plasmide pBIOC32 a été digéré doublement par SacI et BamHI pour supprimer la séquence codant pour le peptide signal naturel de l' $\alpha$ 1-antitrypsine humaine. Cette séquence a été remplacée par celle codant pour le peptide signal PS de 23 acides aminés (ATG AAA GCC TTC ACA CTC GCT CTC TTC TTA GCT CTT TCC

CTC TAT CTC CTG CCC AAT CCA GCC CAT TCC) fusionnée aux 3 premiers codons de la protéine AAT mature (Glu-Asp-Pro). La séquence "PS - 3 premiers codons de la protéine AAT mature" a été amplifiée par PCR à partir du plasmide pMAT103 (Matuoka et Nakamura, 1991) à l'aide des 2 oligodésoxynucléotides, 5' gcgagctcgaattcaacaATG AAA GCC TTC ACA CTC GC 3' (contenant les sites EcoRI et SacI et le signal traductionnel de Kozack) et 5' CT GGG ATC CTC GGA ATG GGC TGG ATT GGG CAG G 3' (contenant le site BamHI de la séquence de la protéine AAT mature), en suivant le protocole d'amplification PCR décrit ci-dessus dans le paragraphe a.1. Après double digestion enzymatique par SacI et BamHI, les fragments d'ADN issus de l'amplification PCR ont été purifiés par électrophorèse sur gel d'agarose 2%, électroélus (Sambrook et al., 1989), précipités en présence de 1/10 de volume de 3M acétate de sodium pH4,8 et de 2,5 volumes d'éthanol absolu à -80°C pendant 30 min., centrifugés à 12000g pendant 30 min., lavés à l'éthanol 70%, séchés, puis ligués à l'ADN plasmidique de pBIOC32 doublement digéré par SacI et BamHI, purifié par électrophorèse sur gel d'agarose 0,8%, électroélué (Sambrook et al., 1989), soumis à la précipitation alcoolique, séché et déphosphorylé par l'enzyme phosphatase alcaline d'intestin de veau (Boehringer Mannheim) selon les recommandations du fabricant. La ligation a été réalisée avec 100 ng du vecteur déphosphorylé décrit ci-dessus et 50 ng de fragments d'ADN digérés issus de l'amplification PCR décrits ci-dessus dans un milieu réactionnel de 10  $\mu$ l en présence de 1  $\mu$ l de tampon T4 DNA ligase x 10 (Amersham) et de 2,5 U d'enzyme T4 DNA ligase (Amersham) à 14°C pendant 16 heures. Les bactéries, *Escherichia coli* DH5 $\alpha$  rendues préalablement compétentes, ont été transformées (Hanahan, 1983).

L'ADN plasmidique des clones obtenus, sélectionnés sur 50 µg/ml ampicilline, a été extrait selon la méthode de la lyse alcaline (Birnboim et Doly, 1979) et analysé par digestion enzymatique par des enzymes de restriction. L'ADN plasmidique de certains des clones retenus a été vérifié par séquençage à l'aide du kit de séquençage T7<sup>TM</sup> commercialisé par Pharmacia selon la méthode des didésoxynucléotides (Sanger et al., 1977). Les séquences du PS et de AAT mature ont été clonées en maintenant leurs phases de lecture ouverte. La séquence de clivage entre les séquences PS et AAT mature est Ser-Glu. Le plasmide résultant a été appelé pBIOC35.

### C. CONSTRUCTION DES PLASMIDES UTILISABLES EN TRANSFORMATION GENETIQUE DU MAIS PAR BIOLISTIQUE

#### C.1. EXPRESSION CONSTITUTIVE : Construction de pPAR-IAR-PA-AAT, pPAR-IAR-PS-AAT, pPAR-IAR-PS-AAT-KDEL et pPAR-IAR-PPS-AAT.

L'expression constitutive dans le maïs a nécessité, entre autre, les séquences régulatrices suivantes :

- promoteur actine de riz suivi de l'intron actine de riz (pAR-IAR) contenu dans le plasmide pAct1-F4 décrit par McElroy et al. (1991) ;
- la séquence terminatrice de transcription, terminateur polyA NOS, qui correspond à la région en 3' non codante du gène de la nopaline synthase du plasmide Ti d'*Agrobacterium tumefaciens* souche à nopaline (Depicker et al., 1982).

Le plasmide pBSII-PAR-IAR-tNOS, dans lequel ont été insérées les séquences codant "PA-AAT", "PS-AAT", "PS-AAT-KDEL" et "PPS-AAT", est décrit par exemple

dans la demande de brevet WO9633277 qui est incorporée par référence.

Les séquences, codant "PA-AAT", "PS-AAT", "PS-AAT-KDEL" et "PPS-AAT", ont été isolées par digestion enzymatique EcoRI respectivement à partir de pBIOC74, pBIOC33, pBIOC35 et pBIOC34. Les fragments digérés ont été purifiés par électrophorèse sur gel d'agarose à 0,8%, puis soumis à l'électroélution, à la précipitation alcoolique, séchés, repris dans H<sub>2</sub>O. Ils ont été traités par action de l'enzyme Klenow (New England Biolabs) selon les recommandations du fabricant. Ils ont été insérés dans le plasmide pBSII-PAR-IAR-tNOS digéré doublement par SalI et NcoI, purifié, traité à l'enzyme Mung Bean Nucléase (New England Biolabs) et déphosphorylé par l'enzyme phosphatase alcaline de veau (Boehringer Mannheim) selon les recommandations des fabricants. La ligation a été réalisée avec 20 ng du vecteur déphosphorylé et 200 ng de fragments d'ADN contenant la séquence codant "PA-AAT", "PS-AAT", "PS-AAT-KDEL" ou "PPS-AAT", décrits ci-dessus, dans un milieu réactionnel de 20 µl en présence de 2 µl de tampon T4 DNA ligase x 10 (Amersham), de 2 µl de 50% polyéthylène glycol 8000 et de 5 U d'enzyme T4 DNA ligase (Amersham) à 14°C pendant 16 heures. Les bactéries, *Escherichia coli* DH5a rendues préalablement compétentes, ont été transformées (Hanahan, 1983). L'ADN plasmidique des clones obtenus, sélectionnés sur 50 µg/ml ampicilline, a été extrait selon la méthode de la lyse alcaline (Birnboim et Doly, 1979) et analysé par digestion enzymatique par des enzymes de restriction. Les plasmides résultants ont été appelés pPAR-IAR-PA-AAT, pPAR-IAR-PS-AAT, pPAR-IAR-PS-AAT-KDEL et pPAR-IAR-PPS-AAT.

## C.2. EXPRESSION DANS L'ALBUMEN DES SEMENCES DE MAIS

### C.2.1 CONSTRUCTION DE pPgzéine-PA-AAT, pPgzéine-PS-AAT, pPgzéine-PS-AAT-KDEL et pPgzéine-PPS-AAT.

L'expression dans l'albumen des semences de maïs a nécessité, entre autre, les séquences régulatrices suivantes :

- le promoteur du gène gamma (g) zéine de maïs (Pgzéine) contenu dans le plasmide p63 est décrit dans Reina et al., 1990. Le plasmide p63 résulte du clonage de Pgzéine, en remplacement du promoteur 35S (P35S), aux sites HindIII et XbaI du plasmide pUC18 renfermant, entre ses sites HindIII et EcoRI, la cassette d'expression "P35S-gus-TNOS" de pBI221 commercialisé par Clontech. Il permet une expression dans l'albumen des semences de maïs.
- la séquence terminatrice de transcription, terminateur polyA NOS, qui correspond à la région en 3' non codante du gène de la nopaline synthase du plasmide Ti d'*Agrobacterium tumefaciens* souche à nopaline (Depicker et al., 1982).

Les plasmides pPgzéine-PA-AAT, pPgzéine-PS-AAT, pPgzéine-PS-AAT-KDEL et pPgzéine-PPS-AAT où les séquences codant "PA-AAT", "PS-AAT", "PS-AAT-KDEL" ou "PPS-AAT" ont été placées sous contrôle du Pgzéine, ont été obtenus par clonage aux sites, SacI et BamHI, traités par l'enzyme T4 DNA polymérase (New England Biolabs), puis déphosphorylés par l'enzyme phosphatase alcaline de veau (Boehringer Mannheim) du plasmide p63, des fragments EcoRI traité à la Klenow (New England Biolabs) isolé à partir de pBIOC74, pBIOC33, pBIOC35 et pBIOC34. La ligation a été réalisée comme

décrite dans l'exemple C1. Les bactéries, *Escherichia coli* DH5a rendues préalablement compétentes, ont été transformées (Hanahan, 1983). L'ADN plasmidique des clones obtenus, sélectionnés sur 50 µg/ml ampicilline, a été extrait selon la méthode de la lyse alcaline et analysé par digestion enzymatique par des enzymes de restriction. Les clones résultants ont été appelés pPgzéine-PA-AAT, pPgzéine-PS-AAT, pPgzéine-PS-AAT-KDEL et pPgzéine-PPS-AAT.

**D. CONSTRUCTION DES PLASMIDES UTILISABLES EN TRANSFORMATION GENETIQUE DU MAIS PAR *Agrobacterium tumefaciens*.**

Les plasmides pSB-PA-AAT, pSB-PS-AAT, pSB-PS-AAT-KDEL, pSB-PPS-AAT, qui portent une répllication de pBR322 et un gène de résistance à la spectinomycine, ont été obtenus par insertion d'un plasmide binaire possédant des régions ori et cos, construit selon les méthodes décrites dans les brevets WO 9400977 et WO 9506722 entre les bordures droite et gauche du T-DNA des cassettes d'expression "promoteur actine de riz-intron actine de riz-gène bar-terminateur NOS" et respectivement PA-AAT sous contrôle de PAR-IAR ou Pgzéine, PS-AAT sous contrôle de PAR-IAR ou Pgzéine, PS-AAT-KDEL sous contrôle de PAR-IAR ou Pgzéine et PPS-AAT sous contrôle de PAR-IAR ou Pgzéine.

Les plasmides obtenus ont été introduits dans *Agrobacterium tumefaciens* souche LBA4404 (pSB1) décrite par Ishida et al. (Nature Biotechnology, 1996, 14 : 745-750). Les bactéries résultantes contiennent les cointégrats de pSB1 et respectivement de pSB-PA-AAT, pSB-PS-AAT, pSB-PS-AAT-KDEL, pSB-PPS-AAT.

## II. OBTENTION DE PLANTES DE MAÏS TRANSGENIQUES.

### A. Obtention et utilisation de cal de maïs comme cible pour la transformation génétique.

La transformation génétique du maïs, quelle que soit la méthode employée (électroporation; *Agrobacterium*, microfibres, canon à particules), requiert généralement l'utilisation de cellules indifférenciées en divisions rapides ayant conservé une aptitude à la régénération de plantes entières. Ce type de cellules compose le cal friable embryogène (dit de type II) de maïs.

Ces cals sont obtenus à partir d'embryons immatures de génotype H1 II ou (A188 x B73) selon la méthode et sur les milieux décrits par Armstrong (1994). Les cals ainsi obtenus sont multipliés et maintenus par repiquages successifs tous les quinze jours sur le milieu d'initiation.

Des plantules sont ensuite régénérées à partir de ces cals en modifiant l'équilibre hormonal et osmotique des cellules selon la méthode décrite par Vain et al. (1989). Ces plantes sont ensuite acclimatées en serre où elles peuvent être croisées ou autofécondées.

### B. Utilisation du canon à particules pour la transformation génétique du maïs.

Le paragraphe précédent décrit l'obtention et la régénération des lignées cellulaires nécessaires à la transformation ; on décrit ici une méthode de transformation génétique conduisant à l'intégration stable des gènes modifiés dans le génome de la plante. Cette méthode repose sur l'utilisation d'un canon à particules ; les cellules cibles sont des fragments de cals décrits dans le paragraphe 1. Ces fragments d'une surface de 10 à 20 mm<sup>2</sup> ont été disposés, 4 h avant

bombardement, à raison de 16 fragments par boîte au centre d'une boîte de petri contenant un milieu de culture identique au milieu d'initiation, additionné de 0,2 M de mannitol + 0,2 M de sorbitol. Les plasmides portant les gènes à introduire sont purifiés sur colonne Qiagen, en suivant les instructions du fabricant. Ils sont ensuite précipités sur des particules de tungsten (M10) en suivant le protocole décrit par Klein (1987). Les particules ainsi enrobées sont projetées vers les cellules cibles à l'aide du canon et selon le protocole décrit par J. Finer (1992).

Les boîtes de cals ainsi bombardées sont ensuite scellées à l'aide de Scellofrais®, puis cultivées à l'obscurité à 27°C. Le premier repiquage a lieu 24 h après, puis tous les quinze jours pendant 3 mois sur milieu identique au milieu d'initiation additionné d'un agent sélectif. Les agents sélectifs utilisables consistent généralement en composés actifs de certains herbicides (Basta®, Round up®,) ou certains antibiotiques (Hygromycine, Kanamycine...).

On obtient après 3 mois ou parfois plus tôt, des cals dont la croissance n'est pas inhibée par l'agent de sélection, habituellement et majoritairement composés de cellules résultant de la division d'une cellule ayant intégré dans son patrimoine génétique une ou plusieurs copies du gène de sélection. La fréquence d'obtention de tels cals est d'environ 0,8 cal par boîte bombardée.

Ces cals sont identifiés, individualisés, amplifiés puis cultivés de façon à régénérer des plantules. Afin d'éviter toute interférence avec des cellules non transformées toutes ces opérations sont menées sur des milieux de culture contenant l'agent sélectif.



Les plantes ainsi régénérées sont acclimatées puis cultivées en serre où elles peuvent être croisées ou autofécondées.

### C. Utilisation d'*Agrobacterium tumefaciens* pour la transformation génétique du maïs.

La technique utilisée est décrite par Ishida et al. (Nature Biotechnology, 1996, 14 : 745-750).

### III. : ANALYSE DE L'EXPRESSION DE L' $\alpha$ 1-ANTITRYPSINE HUMAINE DANS LES GRAINES DE MAIS :

Le protocole d'extraction de l' $\alpha$ 1-antitrypsine pour la réalisation des tests ELISA et des Western blot, à partir de graines de maïs est le suivant : 100 mg de graines sont broyées dans l'azote liquide puis dans 1 ml de tampon Tris-HCl 100mM pH 8 additionné d'EDTA 1mM, de dithiotréitol 1mM et de NaCl 250 mM. Le broyat est conservé à 4°C pour macération pendant 2 heures, puis centrifugé à 4°C pendant 10 min à 10000g.

Sur le surnageant de centrifugation, un dosage des protéines solubles est réalisé par la méthode de Bradford (1976).

#### Immunodétection par Western blot

Les protéines extraites selon le protocole ci-dessus sont dénaturées par chauffage à 95°C pendant 5 min en présence de tampon Tris-HCl 50 mM pH6,8, SDS 4%,  $\beta$ -mercaptoéthanol 1%, saccharose 20% et bleu de bromophénol 0,01%. Les protéines sont ensuite séparées par électrophorèse sur gel de polyacrylamide en conditions dénaturantes selon la technique de Laemmli (1970) à raison de 30  $\mu$ g de protéines solubles par

échantillon. Après migration, les protéines sont transférées sur une membrane de nitrocellulose. Un anticorps polyclonal anti- $\alpha_1$ -antitrypsine obtenu chez le lapin (Dako) est utilisé comme sonde et la révélation est effectuée au moyen d'un anticorps anti-immunoglobulines de lapin marqué à la phosphatase alcaline (Sigma).

Purification partielle de l' $\alpha_1$ -antitrypsine à partir de graines de maïs :

L'extrait protéique est passé sur colonne d'héparine BIORAD, colonne échangeuse d'anions Mono Q, et une colonne gel filtration Superdex 75.

## REFERENCES BIBLIOGRAPHIQUES :

An G., 1986, Plant Physiol., 81, 86-91.

Berg D.E. et Berg C.M., 1983, Biotechnology, 1, 417-435.

Birnboim H.C. et Doly J., 1979, Nucl. Acids. Res., 7, 1513-1523.

Bradford M., 1976, Anal. Biochem., 72, 248.

Close J.J. et Rodriguez R.L., 1982, Gene, 20, 305-316.

Depicker et al., 1982, J. Mol. Appl. Genet., 1, 561-573.

Depigny-This et al., 1992, Plant. Mol. Biol., 20, 467-479.

Franck et al., 1980, Cell, 21, 285-294.

Gamborg O.L., Miller R.A. et Ojima K., 1968, Exp. Cell Res., 50, 151-158.

Guerineau F. et Mullineaux P., 1993, Plant Molecular Biology LABFAX, Groy R.R.D (Ed), Nios. Scientific Publishers, Blackwell Scientific Publications, 121-147.

Hanahan D., 1983, J. Mol. Biol., 166, 557-580.

Holsters et al., 1978, Mol. Gen. Genet., 163, 181-187.

Horsch R.B., Fry J.E., Hoffmann N.L., Eichholtz D. et Rogers S.G., 1985, Science, 227, 1229-1231.

Kay et al., 1987, Science, 236, 1299-1302.

Laemmli U.K., 1970, Nature, 227, 680-685.

Liu et al., 1993, Mol. Plant Microb. Interactions, 6, 144-156.

Matsuoka K. et Nakamura K., 1991, Proc. Natl. Acad. Sci., USA, 88, 834-838.

Montreuil J. 1980, Advances in Carbohydrate Chemistry and Biochemistry, Academic Press Inc., 37, 157-223.

Murakami et al., 1986, Plant Mol. Biol., 7, 343-355.

Murashige T. et Skoog F., 1962, Physiol. Plantarum, 15, 473-497.

Renart J. et Sandoval IV., 1984, Meth. Enzymol., 104, 455-460.

Sambrook et al., 1989, Molecular Cloning Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press.

Sanger et al., 1977, Proc. Natl. Acad. Sci., 74, 5463-5467.

Travis J. et Johnson D., 1981, Meth. in Enzymol., 80, 754-765.

Vliegenthart J.F.G. and Montreuil J., Chap. II, Glycoproteins, Montreuil J., Schachter H., vliegenthart J.F.G., 1995, ELSEVIER SCIENCE B.V., 13-28.

## REVENDEICATIONS

1 - Procédé de production d' $\alpha_1$ -antitrypsine ( $\alpha_1$ -AT), ou de variantes de celle-ci caractérisé par :

- i) l'introduction dans des cellules végétales monocotylédones, d'un acide nucléique comportant une séquence codant pour l' $\alpha_1$ -antitrypsine ( $\alpha_1$ -AT) ou pour une variante de celle-ci, les variantes se différenciant de l' $\alpha_1$ -AT humaine naturelle par une ou plusieurs substitution(s), délétion(s) ou insertion(s) d'acides aminés, et éventuellement une séquence codant pour un agent sélectif ;
- ii) sélection des cellules ayant intégré de manière stable l'acide nucléique ;
- iii) propagation des cellules transformées en culture, ou régénération de plantes entières chimériques ou transgéniques ;
- iv) récupération et éventuellement purification de l' $\alpha_1$ -antitrypsine ou de ses variantes, ainsi produites.

2 - Procédé selon la revendication 1 caractérisé en ce que l' $\alpha_1$ -antitrypsine ou ses variantes présentent une activité d'inhibition de protéases à sérine, de préférence une activité de l' $\alpha_1$ -AT humaine.

3 - Procédé selon la revendication 2 caractérisé en ce que l'acide nucléique introduit dans les cellules végétales code pour la protéine illustrée dans la figure 1, ou pour une protéine présentant au moins 90% d'homologie avec celle-ci, ou encore un fragment de l'une de ces protéines ayant une longueur comprise entre 200 et 393 acides aminés.

4 - Procédé selon la revendication 3 caractérisé en ce que l'acide nucléique comporte en outre une séquence codant pour un peptide signal ou "prépeptide", et éventuellement une séquence codant

pour un signal de rétention endoplasmique ou pour un signal d'adressage vacuolaire.

5 - Procédé selon l'une quelconque des revendications 3 ou 4 caractérisé en ce que l'acide nucléique comporte en outre des séquences régulatrices de transcription reconnues par une cellule végétale monocotylédone.

6 - Protéine présentant une activité d'inhibiteur de protéase à sérine, de préférence une activité de l' $\alpha_1$ -antitrypsine humaine, caractérisée en ce qu'elle est susceptible d'être produite par le procédé selon l'une quelconque des revendications 1 à 5.

7 - Protéine selon la revendication 6 caractérisée en ce qu'elle est N-glycosylée par au moins un glycanne de type mannosidique ou polymannosidique, et/ou par au moins un glycanne de type complexe comportant éventuellement au sein de sa structure un ou plusieurs résidus de xylose et éventuellement un ou plusieurs résidus de fucose.

8 - Protéine selon la revendication 7 comportant la séquence en acides aminés illustrée dans la figure 1, ou une séquence présentant au moins 90% d'homologie avec cette séquence, ou encore un fragment de l'une de ces séquences ayant une longueur comprise entre 200 et 393 acides aminés.

9 - Protéine selon la revendication 8 caractérisée en ce que le glycanne est lié N-glycosidiquement à Asn-46, Asn-83 ou Asn-247.

10 - Protéine selon l'une quelconque des revendications précédentes caractérisée en ce qu'elle porte à la fois un ou plusieurs glycannes de type polymannosique et un ou plusieurs glycannes de type complexe.

11 - Protéine selon l'une quelconque des revendications précédentes caractérisée en ce que le

glycane polymannosique présente la structure GlcNAc<sub>2</sub>-Man<sub>5-9</sub>, par exemple GlcNAc<sub>2</sub>-Man<sub>9</sub>.

12 - Protéine selon l'une quelconque des revendications précédentes caractérisée en ce que le glycane complexe présente la structure GlcNAc<sub>2</sub> (Fuc) Man<sub>3</sub> (Xyl).

13 - Protéine selon l'une des revendications 6 à 12 caractérisée en ce que la séquence en acides aminés comporte en outre un peptide signal N-terminal, permettant une sécrétion de la protéine et éventuellement un signal responsable de la rétention de la protéine dans le réticulum endoplasmique, ou un signal reconnu par une cellule végétale comme étant un signal d'adressage vacuolaire.

14 - Protéine selon la revendication 13 caractérisée en ce que le peptide signal est le prépeptide de la sporamine, de la lectine d'orge, de l'extensine végétale, de l' $\alpha$ -mating factor, des protéines de pathogénèse.

15 - Protéine selon la revendication 13 caractérisée en ce que le peptide signal est celui de l' $\alpha_1$ -antitrypsine native.

16 - Protéine selon la revendication 13 caractérisée en ce que le signal responsable de la rétention de la protéine dans le réticulum endoplasmique est choisi parmi les peptides KDEL, SEKDEL et HDEL.

17 - Protéine selon la revendication 13 caractérisée en ce que le signal d'adressage vacuolaire est d'origine végétale, tel que celui de la sporamine, ou celui de la lectine d'orge.

18 - Protéine selon la revendication 6 caractérisée en ce qu'elle est exempte de chaînes N-glycosidiques.

19 - Acide nucléique comportant : i) une séquence codant pour l' $\alpha_1$ -AT ou pour une variante de celle-ci,

et ii) une séquence d'adressage choisie parmi une séquence codant pour un peptide signal "pré" d'origine végétale et éventuellement une séquence codant pour un signal de rétention ou pour un signal d'adressage vacuolaire, et iii) des séquences régulatrices de transcription reconnues par une cellule végétale monocotylédone.

20 - Acide nucléique selon la revendication 19 caractérisé en ce que les séquences régulatrices comprennent un ou plusieurs promoteur(s) d'origine végétale ou virale.

21 - Acide nucléique selon l'une quelconque des revendications 19 et 20 caractérisé en ce que la séquence codant pour l' $\alpha_1$ -AT est un ADNc codant pour l' $\alpha_1$ -AT mature.

22 - Vecteur comprenant un acide nucléique selon l'une quelconque des revendications 19 à 21.

23 - Cellules végétales monocotylédones transformées de manière stable par un acide nucléique selon l'une quelconque des revendications 19 à 21, et capable de produire une ou plusieurs protéine(s) ayant une activité d'inhibiteur de protéase à sérine, de préférence une activité d' $\alpha_1$ -antitrypsine.

24 - Cellules végétales monocotylédones selon la revendication 23 caractérisées en ce qu'il s'agit d'une culture de cellules végétales, par exemple en milieu liquide ou immobilisées, ou une culture de racines.

25 - Cellules végétales selon la revendication 24 caractérisées en ce qu'il s'agit de cellules faisant partie d'une plante monocotylédone transformée.

26 - Plante monocotylédone chimérique ou transgénique capable de produire une protéine ayant une activité d'inhibiteur de protéase à sérine, de préférence une activité d' $\alpha_1$ -antitrypsine, caractérisée



en ce qu'elle comporte des cellules selon la revendication 23.

27 - Graines de plante transgénique selon la revendication 26.

28 - Anticorps monoclonaux ou polyclonaux capable de reconnaître spécifiquement la protéine selon les revendications 6 à 18.

29 - Produit pharmaceutique comprenant une ou plusieurs protéine(s) selon l'une quelconque des revendications 6 à 18 en association avec un excipient acceptable du point de vue physiologique.

30 - Protéine ayant une activité d' $\alpha_1$ -AT selon les revendications 6 à 18 pour l'utilisation comme médicament.

31 - Protéine selon la revendication 30 pour une utilisation en thérapie de conditions liées à une déficience en  $\alpha_1$ -AT, notamment l'emphysème pulmonaire, la mucoviscidose, le choc septique, ou comme anti-rhumatoïde.

32 - Utilisation d'une protéine selon l'une quelconque des revendications 6 à 18 pour la préparation d'un médicament pour le traitement de conditions liées à une déficience en  $\alpha_1$ -AT.

33 - Utilisation d'une protéine selon l'une quelconque des revendications 6 à 18 dans la préparation de produits cosmétiques, ou de réactifs chimiques.

34 - Extrait de plante contenant 1 à 90% environ d'une protéine selon l'une quelconque des revendications 6 à 18, et notamment l' $\alpha_1$ -antitrypsine.

35 - Extrait de plante selon la revendication 34 pour l'utilisation comme médicament.

36- Extrait de plante selon la revendication 35 pour une utilisation en thérapie de conditions liées à une déficience en  $\alpha_1$ -AT, notamment l'emphysème

pulmonaire, la mucoviscidose, le choc septique ou comme anti-rhumatoïde.

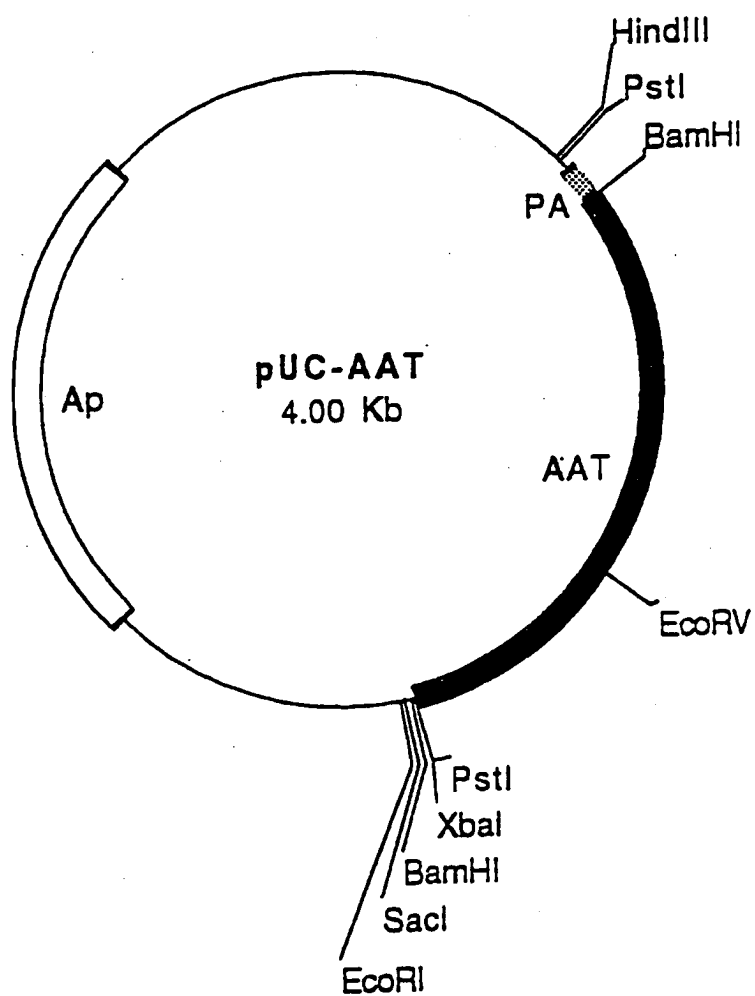
37- Composition pharmaceutique comprenant un extrait de plante selon la revendication 34 en association avec un excipient acceptable du point de vue physiologique, de préférence pour l'administration par voie orale ou nasale, par exemple sous forme d'aérosol.

5' GCGCGGCGGCGCGG CA CCA CCA CTC ACC  
 10 20  
 -1  
 TCG GAC AAT GAA TCG ACA ATG CCG TCT TCT CTC TCG TCG GCG ATC CTC CTC GCG CCA GGC CTC GTC CCT CTC TCC CTC GCT  
 30 40 50 60 70 80 90 100 110  
 -10  
 Met Pro Ser Ser Val Ser Trp Gly Ile Leu Leu Leu Ala Gly Leu Cys Cys Leu Val Pro Val Ser Leu Ala  
 120 130 140 150 160 170 180 190 200  
 +1  
 Glu Asp Pro Gln Gly Asp Ala Ala Gln Lys Thr Asp Thr Ser His Asp Cln Asp His Pro Thr Phe Asn Lys Ile Thr Pro Asn Leu  
 CAG GAT CCC CAG CCA GAT GCT GCG CAG AAG ACA GAT ACA TCC CAC CAT GAT CAG CAT CAC CCA ACC TTC AAC AAG ATC ACC CCC AAC CTC  
 210 220 230 240 250 260 270 280 290  
 60  
 Ala Glu Phe Ala Phe Ser Leu Tyr Arg Gln Leu Ala His Gln Ser Asn Ser Thr Asn Ile Phe Phe Ser Pro Val Ser Ile Ala Thr Ala  
 GCT CAG TTC CCC TTC ACC CTA TAC CCC CAG CTC GCA CAC CAC TCC AAC AGC ACC AAT ATC TTC TTC TCC CCA CTC AGC ATC GCT ACA GCG  
 300 310 320 330 340 350 360 370 380  
 90  
 Phe Ala Met Leu Ser Leu Gly Thr Lys Ala Asp Thr His Asp Glu Ile Leu Glu Gly Leu Asn Phe Asn Leu Thr Glu Ile Pro Glu Ala  
 TTT GCA ATC CTC TCC TCC CTC GCG ACC AAG GCT GAC ACT CAC GAT GAA ATC CTC GAG GCG CTC AAT TTC AAC CTC ACC GAG ATT CCG GAG GCT  
 390 400 410 420 430 440 450 460 470  
 120  
 Gln Ile His Glu Gly Phe Gln Glu Leu Leu Arg Thr Leu Asn Cln Pro Asp Ser Gln Leu Gln Thr Thr Gly Asn Gly Leu Phe Leu  
 CAG ATC CAT GAA GCG TTC CAG GAA CTC CTC CCG TCC ACC CTC ACC GAG CAG CTC CAG CTC ACC ACC GCG AAT GCG CTC TTC CTC  
 480 490 500 510 520 530 540 550 560  
 150  
 Ser Glu Gly Leu Lys Leu Val Asp Lys Phe Leu Glu Asp Val Lys Lys Leu Tyr His Ser Glu Ala Phe Thr Val Asn Phe Gly Asp Thr  
 AGC CAG GCG CTC AAG CTA CTC GAT AAG TTT TTC GAG GAT GTT AAA AAG TTG TAC CAC TCA GAA GCC TTC ACT GTC AAC TTC CCG GAC ACC  
 570 580 590 600 610 620 630 640 650  
 180  
 Glu Glu Ala Lys Lys Gln Ile Asn Asp Tyr Val Glu Lys Gly Thr Gln Gly Lys Ile Val Asp Leu Val Lys Glu Leu Asp Arg Asp Thr  
 GAA GAG GCC AAG AAA CAG ATC AAC CAT TAC CTC CAG AAG CCG ACT CAA GCG AAA ATT GTC GAT TTG CTC AAG GAG GTT GAC ACA GAC ACA  
 660 670 680 690 700 710 720 730 740  
 210  
 Val Phe Ala Leu Val Asn Tyr Ile Phe Phe Lys Gly Lys Trp Glu Arg Pro Phe Glu Val Lys Asp Thr Glu Glu Asp Phe His Val  
 GTT TTT GCT CTC CTC AAT TAC ATC TTC TTT AAA GCC AAA TCG GAG AGA CCC TTT GAA GTC AAG CAC ACC GAG GAA GAG GAC TTC CAC CTC  
 750 760 770 780 790 800 810 820 830

- FIGURE 1A -

220  
 Asp Gln Val Thr Thr Val Lys Val Pro Met Met Lys Arg Leu Gly Met Phe Asn Ile Cln His Cys Lys Lys Leu Ser Ser Trp Val Leu  
 CAC CAG CTC ACC ACC GTG AAG GTG CCT ATG ATG AAG CCT TTA GCC ATG TTT AAC ATC CAG CAC TGT AAG AAG CTC TCC ACC TCG GTG CTC  
 750 760 770 780 790 800 810 820 830 240  
 Leu Met Lys Tyr Leu Gly Asn Ala Thr Ala Ile Phe Phe Leu Pro Asp Glu Gly Lys Leu Gln His Leu Glu Asn Glu Leu Thr His Asp  
 CTG ATG AAA TAC CTG GCC AAT GCC ACC CCT ATC TTT CTC CCT GAT GAG CGG AAA CTA CAG CAC CTC GAA AAT GAA CTC ACC CAC CAC  
 840 850 860 870 880 890 900 910 920 270  
 Ile Ile Thr Lys Phe Leu Glu Asn Glu Asp Arg Arg Ser Ala Ser Leu His Leu Pro Lys Lys Leu Ser Ile Thr Gly Thr Tyr Asp Leu Lys  
 ATC ATC ACC AAG TTC CTG GAA AAT GAA CAC AGA AGG TCT GCC AGC TTA CAT TTA CCC AAA CTG TCC ATT ACT GCA ACC TAT GAT CTG AAG  
 930 940 950 960 970 980 990 1000 1010 300  
 Ser Val Leu Gly Gln Leu Gly Ile Thr Lys Val Phe Ser Asn Gly Ala Asp Leu Ser Gly Val Thr Glu Glu Ala Pro Leu Lys Leu Ser  
 ACC GTG CTG GGT CAA CTG GCC ATC ACT AAG GTG TTA AGC AAT GGG CCT CAC CTC TCC GCG CTC ACA GAG GAG CCA CCC CTC AAG CTC TCC  
 1020 1030 1040 1050 1060 1070 1080 1090 1100 330  
 Lys Ala Val His Lys Ala Val Leu Thr Ile Asp Glu Lys Gly Thr Glu Ala Ala Gly Ala Met Phe Leu Glu Ala Ile Pro Met Ser Ile  
 AAC GCC CTG CAT AAG CCT CTG CTG ACC ATC GAC GAG AAA GGG ACT GAA GCT GCT GGG GCC ATG TTT TTA CAG GCC ATA CCC ATG TCT ATC  
 1110 1120 1130 1140 1150 1160 1170 1180 1190 360  
 Pro Pro Glu Val Lys Phe Asn Lys Pro Phe Val Phe Leu Met Ile Glu Gln Asn Thr Lys Ser Pro Leu Phe Met Gly Lys Val Val Asn  
 CCC CCC GAG GTC AAG TTC AAC AAA CCC TTT GTC TTC TTA ATG ATT GAA CAA AAT ACC AAG TCT CCC CTC TTC ATG GGA AAA CTG GTG AAT  
 1200 1210 1220 1230 1240 1250 1260 1270 1280  
 394  
 Pro Thr Gln Lys STOP  
 CCC ACC CAA AAA TAA CTG CCT CTC GCT CCT CAA CCC CTC CCC TCC ATC CCT GGC CCC CTC CCT GGA TGA CAT TAA AGA AGG GTT GAG CTC  
 1290 1300 1310 1320 1330 1340 1350 1360 1370  
 G AAAAAAAAAAAAAA CC 3'  
 1380 1390 1400 1410 1420 1430

- FIGURE 1B -



- FIGURE 2 -

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 99/00195

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/82 C12N15/15 C07K14/81 C12N5/10 A01H5/00  
A01H5/10 C07K16/38 A61K38/57 A61K7/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K C12N A01H A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	TERASHIMA M ET AL: "PRODUCTION OF FUNCTIONAL HUMAN ALPHA1-ANTITRYPSIN BY RICE CELL CULTURE" ABSTRACTS OF PAPERS. ACS NATIONAL MEETING, 7 September 1997, page COMPLETE 1 XP002069835	1-13, 15, 19-28
Y	see the whole document	4, 13-17, 19, 22-37
Y	WO 96 33277 A (BIOCEM S A ; INST RECHJOVEINAL (FR); LENEÉ PHILIPPE (FR); GRUBER VE) 24 October 1996 cited in the application see page 3, line 1 - page 4, line 26 see page 8, line 5 - page 9, line 20 -/--	4, 13-17, 19, 22-37



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

### \* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

9 June 1999

Date of mailing of the international search report

15/06/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Van der Schaal, C

# INTERNATIONAL SEARCH REPORT

Intern. Application No

PCT/FR 99/00195

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 137 633 A (ZYMOGENETICS INC) 17 April 1985	6, 18
Y	see the whole document	29-33, 35-37
P, X	WO 98 36085 A (APPLIED PHYTOLOGICS INC ; RODRIGUEZ RAYMOND L (US); SUTLIFF THOMAS) 20 August 1998	1-13, 15, 19-28
Y	see claims; examples 1, 2	4, 13-17, 19, 22-37

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/FR 99/00195

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9633277 A	24-10-1996	FR 2733249 A	25-10-1996
		AU 5696796 A	07-11-1996
		BG 101959 A	31-07-1998
		BR 9608011 A	05-01-1999
		CA 2218418 A	24-10-1996
		CN 1185178 A	17-06-1998
		CZ 9703309 A	18-03-1998
		EP 0822988 A	11-02-1998
		HU 9801438 A	28-10-1998
		PL 322874 A	02-03-1998
		SK 140197 A	08-07-1998
EP 0137633 A	17-04-1985	AT 70302 T	15-12-1991
		AU 588793 B	28-09-1989
		AU 3180184 A	14-02-1985
		DE 3485336 A	23-01-1992
		JP 60186290 A	21-09-1985
WO 9836085 A	20-08-1998	AU 6171698 A	08-09-1998



# RAPPORT DE RECHERCHE INTERNATIONALE

Dem. e internationale No

PCT/FR 99/00195

## A. CLASSEMENT DE L'OBJET DE LA DEMANDE

CIB 6 C12N15/82 C12N15/15 C07K14/81 C12N5/10 A01H5/00  
A01H5/10 C07K16/38 A61K38/57 A61K7/00

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

## B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 6 C07K C12N A01H A61K

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

## C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	TERASHIMA M ET AL: "PRODUCTION OF FUNCTIONAL HUMAN ALPHA1-ANTITRYPSIN BY RICE CELL CULTURE" ABSTRACTS OF PAPERS. ACS NATIONAL MEETING, 7 septembre 1997, page COMPLETE 1 XP002069835	1-13, 15, 19-28
Y	voir le document en entier	4, 13-17, 19, 22-37
Y	WO 96 33277 A (BIOCEM S A ; INST RECHJOVEINAL (FR); LENEÉ PHILIPPE (FR); GRUBER VE) 24 octobre 1996 cité dans la demande voir page 3, ligne 1 - page 4, ligne 26 voir page 8, ligne 5 - page 9, ligne 20 --- -/-	4, 13-17, 19, 22-37

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

### \* Catégories spéciales de documents cités:

- "A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- "E" document antérieur, mais publié à la date de dépôt international ou après cette date
- "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

9 juin 1999

Date d'expédition du présent rapport de recherche internationale

15/06/1999

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Van der Schaal, C

# RAPPORT DE RECHERCHE INTERNATIONALE

Dem. : Internationale No

PCT/FR 99/00195

## C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	EP 0 137 633 A (ZYMOGENETICS INC) 17 avril 1985	6,18
Y	voir le document en entier	29-33, 35-37
	----	
P,X	WO 98 36085 A (APPLIED PHYTOLOGICS INC ; RODRIGUEZ RAYMOND L (US); SUTLIFF THOMAS) 20 août 1998	1-13,15, 19-28
Y	voir revendications; exemples 1,2	4,13-17, 19,22-37
	-----	

# RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Dem. internationale No

PCT/FR 99/00195

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9633277 A	24-10-1996	FR 2733249 A	25-10-1996
		AU 5696796 A	07-11-1996
		BG 101959 A	31-07-1998
		BR 9608011 A	05-01-1999
		CA 2218418 A	24-10-1996
		CN 1185178 A	17-06-1998
		CZ 9703309 A	18-03-1998
		EP 0822988 A	11-02-1998
		HU 9801438 A	28-10-1998
		PL 322874 A	02-03-1998
		SK 140197 A	08-07-1998
EP 0137633 A	17-04-1985	AT 70302 T	15-12-1991
		AU 588793 B	28-09-1989
		AU 3180184 A	14-02-1985
		DE 3485336 A	23-01-1992
		JP 60186290 A	21-09-1985
WO 9836085 A	20-08-1998	AU 6171698 A	08-09-1998